

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 15:54:17 ; Search time 60 seconds  
(without alignments)  
1196.118 Million cell updates/sec

Title: US-09-895-814-525  
Perfect score: 1369  
Sequence: 1 MATAGNPMWFLGYLILGVA.....GVVTNLCCKTEWTEKTVQAS 254

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1369	100.0	254	3	AAB21294 Human KLK
2	1369	100.0	254	4	AAM01174 Human pro
3	1369	100.0	254	4	Aau69819 Human pro
4	1369	100.0	254	4	AAG99059 Human pro
5	1369	100.0	254	4	ABU71710 Prostate
6	1369	100.0	254	5	ABb95279 Human P70
7	1369	100.0	254	6	ABP54360 Human KLK
8	1369	100.0	254	6	ABP54357 Human KLK
9	1369	100.0	254	6	ABR54391 Prostate
10	1369	100.0	254	7	ABD13975 Human pro
11	1369	100.0	1079	4	ABb74830 Prostate
12	1369	100.0	1079	4	ABU71860 Prostate
13	1364	99.6	254	3	AAB21320 Human pro
14	1364	99.6	254	4	Aay72525 Human pro
15	1364	99.6	254	5	AAU74901 Protein s
16	1364	99.6	254	5	AAU74932 Amino aci
17	1364	98.8	254	4	AAM01173 Human pro
18	1352	98.8	254	4	Aau69818 Human pro
19	1352	98.8	254	4	AG99058 Human pro
20	1352	98.8	254	4	ABU71709 Prostate
21	1352	98.8	254	5	ABb95278 Human P70
22	1352	98.8	254	5	ABR54390 Prostate
23	1352	98.8	254	7	ABD13973 Human mat
24	1352	98.8	258	3	AAB21324 Human EMS
25	1342	98.0	249	3	AAB21307 Human pro

## ALIGNMENTS

RESULT 1  
AAB21294  
ID AAB21294 standard; protein; 254 AA.  
XX  
AC AAB21294;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KLK-L1 protein #2.  
XX  
KW Human, KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
KW kallikrein-like protein; serine protease; cytosolic; cancer;  
KW prostate cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-CA000258.  
XX  
PR 11-MAR-1999; 99US-0124260P.  
PR 01-APR-1999; 99US-0127386P.  
PR 21-JUL-1999; 99US-0144919P.  
(MOUN ) MOUNT SINAI HOSPITAL.  
PI Yousef GM, Diamandis EP;  
XX  
DR WPI; 2000-587440/55.  
DR N-PSDB; AAA95896.  
XX  
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
protein mediated disorders, especially cancer.  
XX  
PS Claim 8; Page 141; 184pp; English.  
XX  
CC The present sequence is kallikrein-like protein KLK-L1. Kallikreins and  
CC kallikrein-like proteins are a subgroup of the serine protease enzyme  
CC family. They catalyze the selective cleavage of specific polypeptide  
CC precursors to release peptides with potent biological activity. Nucleic  
CC acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4,  
CC KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the  
CC treatment, monitoring and diagnosis of cancers, especially prostate  
CC cancer. They can also be used to identify a substance that can associate  
CC with or mediate the biological activity of the proteins. Antibodies can  
CC be used to treat conditions mediated by the kallikrein-like proteins

26 1330 97.2 253 3 AAB21308 Human EMS  
27 1255 91.7 234 4 AAE00397  
28 1241 90.7 231 2 AAY25510 Human pro  
29 1237 90.4 237 3 AAB21293 Human KLK  
30 1205 88.0 449 4 AAM01227  
31 1205 88.0 449 4 Aau69872 Human pro  
32 1205 88.0 449 4 ABU71763 Prostate  
33 1205 88.0 449 5 ABB95332  
34 1205 88.0 449 6 ABR54444 Prostate  
35 1205 88.0 449 7 ABD14067 Human pro  
36 1205 88.0 585 4 ABU71889 Prostate  
37 1205 88.0 585 6 ABR54580 Prostate  
38 1205 88.0 585 7 ADB14470 FOPP/hpAP  
39 1205 88.0 585 7 ABU71890 Prostate  
40 1205 88.0 801 4 AAY72526 Human pro  
41 1191 87.0 226 4 AAY72526 Protein s  
42 1191 87.0 226 5 AAU74932 Amino aci  
43 1191 87.0 226 5 AAU74933  
44 1191 87.0 312 4 AAY72522 NS1-P703P  
45 1191 87.0 312 5 AAU74768 Amino aci

XX SQ Sequence 254 AA;  
Query Match 100.0%; Score 1369; DB 3; Length 254;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60  
DB 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 120  
DB 61 VHPQWVLSAAHCFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 120  
QY 121 KLDSESVESDTRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
DB 121 KLDSESVESDTRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
QY 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQGVGVYTNL 240  
DB 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQGVGVYTNL 240  
QY 241 CKFTWIEKTVQAS 254  
DB 241 CKFTWIEKTVQAS 254  
RESULT 2  
AAU01174  
ID AAU01174 standard; protein; 254 AA.  
XX AAU01174;  
XX  
XX 04-OCT-2001 (first entry)  
XX Human prostate-specific amino acid sequence P703P.  
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
XX cytostatic; gene therapy; metastasis.  
XX Homo sapiens.  
XX WO200151633-A2.  
XX 19-JUL-2001.  
XX 16-JAN-2001; 2001WO-US001574.  
XX 14-JAN-2000; 2000US-00483672.  
XX (CORI-) CORIXA CORP.  
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
XX Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;  
XX Wang A, Meagher WJ;  
XX WPI; 2001-425873/45.  
XX New polynucleotide encoding a prostate-specific protein, for diagnosing,  
XX monitoring and treating prostate cancer in a patient and for use in  
XX vaccines.  
XX  
XX Claim 2; Page 405; 543pp; English.  
XX  
XX The present invention describes polynucleotide sequences (I) which encode  
XX prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
XX and can be used in vaccine production and gene therapy. (I), (II),  
XX antibodies to (II), fusion proteins comprising (II), and isolated T cells  
XX prepared using (I) or (II) are used to treat cancer in a patient. (I) and  
XX the antibodies are also used in the detection of cancer in a patient. The  
XX cancer that is diagnosed or treated is particularly prostate cancer. (I)  
XX and (II) can be used in vaccines. The antibodies or (I) can be used for

CC monitoring the progression of cancer in a patient. (I) and (II) can also  
CC be used to improve diagnostic and therapeutic methods for prostate  
CC cancer. They can indicate the level of metastasis as well as the prostate  
CC volume. AAH93357 to AAH93944 and AA01115 to AA01118 represent  
CC polynucleotide and amino acid sequences used in the exemplification of  
CC the present invention  
XX  
XX SQ Sequence 254 AA;  
Query Match 100.0%; Score 1369; DB 4; Length 254;  
Best Local Similarity 100.0%; Pred. No. 6e-96; 0; Indels 0; Gaps 0;  
Matches 254; Conservative 0; Mismatches 0;  
QY 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60  
DB 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 120  
DB 61 VHPQWVLSAAHCFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 120  
QY 121 KLDSESVESDTRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
DB 121 KLDSESVESDTRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
QY 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQGVGVYTNL 240  
DB 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQGVGVYTNL 240  
QY 241 CKFTWIEKTVQAS 254  
DB 241 CKFTWIEKTVQAS 254  
RESULT 3  
AAU69819  
ID AAU69819 standard; protein; 254 AA.  
XX AAU69819;  
XX 30-JAN-2002 (first entry)  
XX Human prostate cDNA encoded protein #27.  
XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.  
XX Homo sapiens.  
XX WO200173032-A2.  
XX 04-OCT-2001.  
XX 27-MAR-2001; 2001WO-US009919.  
XX 27-MAR-2000; 2000US-00536857.  
XX 09-MAY-2000; 2000US-00568100.  
XX 12-MAY-2000; 2000US-00570737.  
XX 13-JUN-2000; 2000US-00593793.  
XX 27-JUN-2000; 2000US-00605783.  
XX 09-AUG-2000; 2000US-00638215.  
XX 29-AUG-2000; 2000US-00651236.  
XX 06-SEP-2000; 2000US-00657279.  
XX 10-OCT-2000; 2000US-00679426.  
XX 09-NOV-2000; 2000US-00709729.  
XX (CORI-) CORIXA CORP.  
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
XX Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
XX Li SX, Wang A, Skeiky YAW, Hepler Wt, Henderson RA;  
XX WPI; 2001-639232/73.

DR N-PSDB; AAS63914.  
 XX New human prostate-specific polypeptides and polynucleotides useful for  
 PT the diagnosis and treatment of cancer, especially prostate cancer.  
 XX  
 XX Claim 2; Page 405-406; 579pp; English.  
 PS  
 XX The invention relates to isolated prostate-specific polynucleotides,  
 CC polypeptides, fusion proteins of the polypeptides, antibodies raised  
 CC against the polypeptides (or antigenic epitopes derived from them) and  
 CC antigen-presenting cells expressing the polypeptides. The antibodies are  
 CC useful for detecting the presence of cancer, especially prostate cancer.  
 CC The polypeptides, polynucleotides and the antigen-presenting cells are  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein, and for inhibiting the development of cancer especially prostate  
 CC cancer. Compositions comprising the polynucleotide and/or polypeptide are  
 CC useful for stimulating an immune response, and for treating cancer. The  
 CC oligonucleotide is useful for detecting cancer. The present sequence is a  
 CC prostate specific polypeptide of the invention  
 XX  
 SQ Sequence 254 AA;  
 Query Match 100.0%; Score 1369; DB 4; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 6e-96;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATAGNFWGFLGYLLILGVAGSLVSGSCQIINGEDCSPHSQPWQAALVMENELFCSGVL 60  
 DB 1 MATAGNFWGFLGYLLILGVAGSLVSGSCQIINGEDCSPHSQPWQAALVMENELFCSGVL 60  
 QY 61 VHPQWLVAACHCFONSYTIGLHLSLEADQEPGQSMVEASLSVRHPEYNRPLLANDLMLI 120  
 DB 61 VHPQWLVAACHCFONSYTIGLHLSLEADQEPGQSMVEASLSVRHPEYNRPLLANDLMLI 120  
 QY 121 KLDESVSSEDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEEVCSK 180  
 DB 121 KLDESVSSEDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEEVCSK 180  
 QY 181 LYDPLVHPSMFCAGGGQDKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240  
 DB 181 LYDPLVHPSMFCAGGGQDKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240  
 QY 241 CKFTWIEKTVQAS 254  
 DB 241 CKFTWIEKTVQAS 254  
 RESULT 4  
 AAG99059  
 ID AAG99059 standard; protein; 254 AA.  
 XX  
 AC AAG99059;  
 XX  
 XX 25-SEP-2001 (first entry)  
 DT Human prostate-specific amino acid of P703P.  
 XX  
 DE Human; prostate cancer; therapy; diagnosis; cat eye syndrome;  
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;  
 KW prostate specific antigen; PSA.  
 XX  
 OS Homo sapiens.  
 XX WO200134802-A2.  
 XX  
 XX 17-MAY-2001.  
 XX  
 XX 09-NOV-2000; 2000WO-US030904.  
 PF  
 XX 12-NOV-1999; 99US-00439313.  
 XX  
 XX 18-NOV-1999; 99US-00443686.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA

XX Xu J, Dillion DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;  
 XX WPI; 2001-308785/32.  
 DR  
 XX Isolated polypeptide comprising at least an immunogenic portion of a  
 PT prostate-specific protein, useful in the diagnosis and therapy of  
 PT prostate cancer.  
 XX  
 PS Claim 3; Page 304-305; 325pp; English.  
 XX  
 CC The present invention describes an isolated polypeptide (P1) comprising  
 CC at least an immunogenic portion of a prostate-specific protein, or its  
 CC variant. Also described are polynucleotides (N1) encoding (P1), (P1) and  
 CC (N1) have cytostatic activity and can be used in vaccine production. The  
 CC polypeptides, nucleic acids and antibodies from the present invention are  
 CC useful in the diagnosis and therapy of prostate cancer. Prostate specific  
 CC genes P704P, P712P, P774P, P775P and B305D are located in a genomic  
 CC region on chromosome 22q11.2 known as the Cat Eye Syndrome region.  
 CC Prostate specific antigen (PSA) P501S was located on chromosome 1.  
 CC AAH84671 to AAH85143 and AAG99000 to AAG99077 represent polynucleotide  
 CC and polypeptide sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 254 AA;  
 Query Match 100.0%; Score 1369; DB 4; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 6e-96;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATAGNFWGFLGYLLILGVAGSLVSGSCQIINGEDCSPHSQPWQAALVMENELFCSGVL 60  
 DB 1 MATAGNFWGFLGYLLILGVAGSLVSGSCQIINGEDCSPHSQPWQAALVMENELFCSGVL 60  
 QY 61 VHPQWLVAACHCFONSYTIGLHLSLEADQEPGQSMVEASLSVRHPEYNRPLLANDLMLI 120  
 DB 61 VHPQWLVAACHCFONSYTIGLHLSLEADQEPGQSMVEASLSVRHPEYNRPLLANDLMLI 120  
 QY 121 KLDESVSSEDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEEVCSK 180  
 DB 121 KLDESVSSEDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEEVCSK 180  
 QY 181 LYDPLVHPSMFCAGGGQDKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240  
 DB 181 LYDPLVHPSMFCAGGGQDKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240  
 QY 241 CKFTWIEKTVQAS 254  
 DB 241 CKFTWIEKTVQAS 254  
 RESULT 5  
 ABU71710  
 ID ABU71710 standard; protein; 254 AA.  
 XX  
 AC ABU71710;  
 XX  
 XX 10-JUN-2003 (first entry)  
 DT Prostate cancer specific antigen P703P #7.  
 XX  
 DE Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 XX Synthetic.  
 OS  
 XX US2002192763-A1.  
 XX  
 XX 19-DEC-2002.  
 PD  
 XX 29-JUN-2001; 2001US-00895793.  
 PF

XX 04-OCT-1999; 99US-0157455P.  
PR 04-OCT-2000; 2000US-00679272.  
PR 28-MAR-2001; 2001US-00822827.  
XX (XUJJ/) XU J.  
PA (DILL/) DILLON D C.  
PA (MITC/) MITCHAM J L.  
PA (HARL/) HARLOCKER S L.  
PA (JIAN/) JIANG Y.  
PA (KALO/) KALOS M D.  
PA (FANG/) FANGER G R.  
PA (RETT/) RETTER M W.  
PA (STOL/) STOLK J A.  
PA (DAYC/) DAY C H.  
PA (VEDV/) VEDVICK T S.  
PA (CART/) CARTER D.  
PA (LISX/) LI S X.  
PA (WANG/) WANG A.  
PA (SKEI/) SKEIKY Y A W.  
PA (HEPL/) HEPLER W T.  
PA (HEND/) HENDERSON R A.  
PA (HURA/) HURAL J.  
PA (MCNE/) MCNEILL P D.  
PA (HOUH/) HOUGHTON R L.  
PA (DBAS/) Y DE BASSOLS C V.  
PA (FOYT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
PI McNeill PD, Houghton RL, Y De Bassols CV, Foy TM;  
XX WPI, 2001-245062/25.

DR Prostate specific protein and its encoding polynucleotide, useful for the  
PT treatment and diagnosis of prostate cancer.

XX Example 3; SEQ ID NO 525; 85pp; English.

PS The invention describes a fusion protein comprising at least one amino  
XX acid sequence of immunogenic portions of any of the 3 sequences not  
CC defined in the specification, or sequences having at least 70 or 90 %  
CC sequence identity to any one of the 35 sequences defined in the USPTO web  
CC site, which is encoded by any of the 4 nucleotide sequences not defined  
CC in the specification. The fusion protein, composition and methods are  
CC useful for diagnosing, preventing and/or treating cancer, particularly  
CC prostate cancer. The proteins are useful as markers to indicate the  
CC presence or absence of cancer. This is the amino acid sequence of a  
CC prostate cancer specific antigen. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the US patent office at  
CC seqdata.uspto.gov/sequence.html?DocID=US20020192763

XX seqdata.uspto.gov/sequence.html?DocID=US20020192763

SQ Sequence 254 AA;

Query Match 100.0%; Score 1369; DB 4; Length 254;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGNPGWFLGYLILGVAGSLVSGSCSIQINGEDCSPHSQPQWQALVMENELFCSGVL 60  
Db 1 MATAGNPGWFLGYLILGVAGSLVSGSCSIQINGEDCSPHSQPQWQALVMENELFCSGVL 60  
QY 61 VHPQWLSAHCQFQNSYITGLHSLQADQEPGQWVEASLSVRHPEYRPLLANDLMLI 120  
Db 61 VHPQWLSAHCQFQNSYITGLHSLQADQEPGQWVEASLSVRHPEYRPLLANDLMLI 120  
QY 121 KLDSESVESDTRISIASQCCTAGNSCLVSGWGLLANGRMPTVLQCVNVSWEVCSK 180  
Db 121 KLDSESVESDTRISIASQCCTAGNSCLVSGWGLLANGRMPTVLQCVNVSWEVCSK 180  
QY 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGFVYTNL 240

Db 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGFVYTNL 240  
QY 241 CKTEWIEKTVQAS 254  
Db 241 CKTEWIEKTVQAS 254  
RESULT 6  
ABB95279  
ID ABB95279 standard; protein; 254 AA.  
XX AC ABB95279;  
XX DT 19-JUL-2002 (first entry)  
XX DE Human P703P putative full length protein SEQ ID NO 525.  
XX KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
XX KW Gene therapy.  
XX OS Homo sapiens.  
XX PN US2002022248-A1.  
XX PD 21-FEB-2002.  
XX PF 12-JAN-2001; 2001US-00759143.  
XX PR 25-FEB-1997; 97US-00806099.  
PR 01-AUG-1997; 97US-00904804.  
PR 10-FEB-1998; 98US-00020956.  
PR 25-FEB-1998; 98US-00030607.  
PR 14-JUL-1998; 98US-00115453.  
PR 23-SEP-1998; 98US-00159812.  
PR 15-JAN-1999; 99US-00232149.  
PR 09-APR-1999; 99US-00288946.  
PR 13-JUL-1999; 99US-00352616.  
PR 12-NOV-1999; 99US-00439313.  
PR 18-NOV-1999; 99US-00433686.  
PR 14-JAN-2000; 2000US-00483672.  
PR 27-MAR-2000; 2000US-00536857.  
PR 09-MAY-2000; 2000US-00568100.  
PR 12-MAY-2000; 2000US-00570737.  
PR 13-JUN-2000; 2000US-00593793.  
PR 27-JUN-2000; 2000US-00605783.  
PR 10-AUG-2000; 2000US-00636235.  
PR 29-AUG-2000; 2000US-00651236.  
PR 05-SEP-2000; 2000US-00657279.  
PR 02-OCT-2000; 2000US-00679426.  
PR 10-OCT-2000; 2000US-00685166.

(XUJJ/) XU J.  
(DILL/) DILLON D C.  
(MITC/) MITCHAM J L.  
(HARL/) HARLOCKER S L.  
(JIAN/) JIANG Y.  
(KALO/) KALOS M D.  
(FANG/) FANGER G R.  
(RETT/) RETTER M W.  
(STOL/) STOLK J A.  
(DAYC/) DAY C H.  
(VEDV/) VEDVICK T S.  
(CART/) CARTER D.  
(LISX/) LI S X.  
(WANG/) WANG A.  
(SKEI/) SKEIKY Y A W.  
(HEPL/) HEPLER W T.  
(HEND/) HENDERSON R A.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2002-255649/30.  
XX New prostate-specific polynucleotides for diagnosing and treating  
XX diseases, in particular prostate cancer, and as markers for the  
PT progression of cancer.  
XX Claim 2; SEQ ID NO 525; 87pp; English.  
XX The present invention provides prostate-specific coding sequences and  
XX their encoded proteins. These can be used in the diagnosis and treatment  
XX of cancers, particularly prostate cancer. The present sequence is a  
XX protein described in the invention  
XX  
SQ Sequence 254 AA;  
Query Match 100.0%; Score 1369; DB 5; Length 254;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATAGNPGWFLGYLILGVAGSLVSGSCQIINGDCSPHSQPWQAALVMENELFCSGVL 60  
Db 1 MATAGNPGWFLGYLILGVAGSLVSGSCQIINGDCSPHSQPWQAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMVEASLSVRHPEYNRPILLANDLMLI 120  
Db 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMVEASLSVRHPEYNRPILLANDLMLI 120  
QY 121 KLDESVSSEDITRISIASQCPTAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
Db 121 KLDESVSSEDITRISIASQCPTAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKSCNGSDGGPLICNGYLGVSFGKAPCGQGVPGVYTNL 240  
Db 181 LYDPLYHPSMFCAGGQDQKSCNGSDGGPLICNGYLGVSFGKAPCGQGVPGVYTNL 240  
QY 241 CKFTEWIEKTVQAS 254  
Db 241 CKFTEWIEKTVQAS 254  
RESULT 7  
ABP54360  
ID ABP54360 standard; protein; 254 AA.  
XX  
XX AC ABP54360;  
XX  
XX DT 20-JAN-2003 (first entry)  
XX Human KLK4 protein SEQ ID NO:13.  
XX  
XX KW Human; KLK4; cancer; benign tumour; cytostatic4.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200277243-A1.  
XX  
XX PD 03-OCT-2002.  
XX  
XX PF 27-MAR-2002; 2002WO-AU0000378.  
XX  
XX PR 27-MAR-2001; 2001AU-00004022.  
XX  
XX PA (UYQU-) UNIV QUEENSLAND TECHNOLOGY.  
XX  
XX PI Dong Y, Clements JA;  
XX  
XX DR WPI; 2003-029939/02.  
XX  
XX DR N-PSDB; ASQ83347.  
XX  
XX Detecting the presence or diagnosing the risk of cancer or benign tumor,  
XX e.g. an ovarian, endometrial or prostate cancer, by determining the  
XX presence of or detecting aberrant expression of KLK4 in a biological

PT sample from the patient.  
XX  
XX Disclosure; Page 118-119; 126pp; English.  
XX The present invention describes a method (M1) for detecting the presence  
XX of diagnosing the risk of cancer or benign tumour in a patient. M1  
XX comprises determining the presence of or detecting aberrant expression of  
XX KLK4 in a biological sample obtained from the patient. KLK4 has  
XX cytostatic activity. The method is useful for detecting the presence or  
XX diagnosing the risk of a cancer or a benign tumour in a patient,  
XX particularly an ovarian, endometrial or prostate cancer, or a cancer or  
XX benign tumour associated with an organ or tissue from the ovaries,  
XX endometrium or prostate. An agent which can be used for restoring or  
XX modulating KLK4 expression can be used for treating or preventing cancer  
XX or benign tumour. KLK4 polynucleotides, polypeptides or antigen-binding  
XX molecules from the present invention can be used for detecting aberrant  
XX KLK4 polynucleotides or aberrant K4 polypeptides that correlate with a  
XX cancer or a benign tumour. The present sequence represents human KLK4  
XX from the present invention  
XX  
SQ Sequence 254 AA;  
Query Match 100.0%; Score 1369; DB 6; Length 254;  
Best Local Similarity 100.0%; Pred. No. 6e-96;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATAGNPGWFLGYLILGVAGSLVSGSCQIINGDCSPHSQPWQAALVMENELFCSGVL 60  
Db 1 MATAGNPGWFLGYLILGVAGSLVSGSCQIINGDCSPHSQPWQAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMVEASLSVRHPEYNRPILLANDLMLI 120  
Db 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMVEASLSVRHPEYNRPILLANDLMLI 120  
QY 121 KLDESVSSEDITRISIASQCPTAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
Db 121 KLDESVSSEDITRISIASQCPTAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKSCNGSDGGPLICNGYLGVSFGKAPCGQGVPGVYTNL 240  
Db 181 LYDPLYHPSMFCAGGQDQKSCNGSDGGPLICNGYLGVSFGKAPCGQGVPGVYTNL 240  
QY 241 CKFTEWIEKTVQAS 254  
Db 241 CKFTEWIEKTVQAS 254  
RESULT 8  
ABP54357  
ID ABP54357 standard; protein; 254 AA.  
XX  
XX AC ABP54357;  
XX  
XX DT 20-JAN-2003 (first entry)  
XX Human KLK4 protein SEQ ID NO:6.  
XX  
XX KW Human; KLK4; cancer; benign tumour; cytostatic.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200277243-A1.  
XX  
XX PD 03-OCT-2002.  
XX  
XX PF 27-MAR-2002; 2002WO-AU0000378.  
XX  
XX PR 27-MAR-2001; 2001AU-00004022.  
XX  
XX PA (UYQU-) UNIV QUEENSLAND TECHNOLOGY.  
XX  
XX PI Dong Y, Clements JA;  
XX  
XX



PN US2003185830-A1.  
XX 02-OCT-2003.  
PD 12-NOV-2002; 2002US-00294025.  
XX 25-FEB-1997; 97US-00806099.  
XX 01-AUG-1997; 97US-00904804.  
XX 09-FEB-1998; 98US-00202056.  
XX 25-FEB-1998; 98US-00030607.  
XX 14-JUL-1998; 98US-00115453.  
XX 23-SEP-1998; 98US-00159812.  
XX 15-JAN-1999; 99US-00232149.  
XX 09-APR-1999; 99US-00288946.  
XX 13-JUL-1999; 99US-00352616.  
XX 12-NOV-1999; 99US-00439313.  
XX 18-NOV-1999; 99US-00436886.  
XX 14-JAN-2000; 2000US-00483672.  
XX 27-MAR-2000; 2000US-00536857.  
XX 09-MAY-2000; 2000US-00568100.  
XX 12-MAY-2000; 2000US-00570737.  
XX 13-JUN-2000; 2000US-00593793.  
XX 27-JUN-2000; 2000US-00605783.  
XX 09-AUG-2000; 2000US-00626215.  
XX 29-AUG-2000; 2000US-00651236.  
XX 06-SEP-2000; 2000US-00657279.  
XX 02-OCT-2000; 2000US-00679426.  
XX 10-OCT-2000; 2000US-00685166.  
XX 09-NOV-2000; 2000US-00709729.  
XX 12-JAN-2001; 2001US-00759143.  
XX 09-FEB-2001; 2001US-00780669.  
XX 09-MAY-2001; 2001US-00852911.  
XX 29-JUN-2001; 2001US-00895814.  
XX 10-DEC-2001; 2001US-00012896.  
XX 09-MAY-2002; 2002US-00146678.  
XX (CORI-) CORIXA CORP.  
XX  
XX Xu J, Stolk JA, Kalos MD;  
XX WPI; 2003-756193/71.  
XX N-PSDB; ADB13974.  
XX  
XX New isolated polypeptide for use in a vaccine for stimulating an immune  
XX response, or for treating or diagnosis cancer, preferably prostate  
XX cancer.  
XX  
XX Example 3; Page; 101pp; English.  
XX  
XX The invention relates to an isolated polypeptide comprising no more than  
XX 11-542 amino acids of ADB13563 comprising a sequence ADB1487. The  
XX peptides comprise a fragment ADB13563 of that contain naturally processed  
XX T-cell epitopes for 3 class I major histocompatibility complex (MHC)  
XX alleles. ADB13563 is a polypeptide encoded by a human prostate specific  
XX cDNA, one of 648 disclosed as new. Also included are nucleic acids  
XX encoding the proteins and peptides, expression vectors, a host cell  
XX transformed with the vector, an isolated antibody (or antigen binding  
XX fragment) that specifically binds to the protein or peptide, detecting  
XX the presence of a cancer in a patient (comprising contacting a patient  
XX sample with a binding agent that binds to the peptides or a polypeptide  
XX appearing as ADB13558, detecting the amount of polypeptide that binds to  
XX the agent and comparing the amount of polypeptide to a predetermined cut-  
XX off value to determine the presence of cancer), a fusion protein  
XX comprising the peptides or proteins, stimulating or expanding T cells  
XX specific for a tumour protein comprising contacting T cells with the  
XX peptides or the isolated T cell population, treating prostate cancer in a  
XX patient comprising administering a composition comprising the peptides,  
XX nucleic acids, antibodies or compounds, determining the presence of a  
XX cancer in a patient and treating prostate cancer in a patient comprising  
XX incubating cluster of differentiation (CD4+) and/or CD8+ T cells isolated  
XX from a patient with the peptides or antigen presenting cells that express  
XX (the peptides so that the T cells proliferate, and administering the  
XX proliferated T cells to the patient. The peptides (or an oligonucleotide

CC that hybridises to nucleic acid encoding them), is used to detect the  
CC presence of cancer in a patient. The peptides, nucleic acids encoding, or  
CC antigen-presenting cells expressing the nucleic acid, are used to  
CC stimulate or expand T cells specific for a tumour protein. The peptides,  
CC nucleic acids, antibodies, fusion proteins, T cell populations or antigen  
CC presenting cells are used to stimulate an immune response or treat  
CC prostate cancer in a patient. The present sequence is a prostate specific  
CC protein of the invention. Note: Except where otherwise indicated, the  
CC sequence data for this patent did not form part of the printed, the  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20030185830.  
XX  
XX Sequence 254 AA;  
XX  
XX Query Match 100.0%; Score 1369; DB 7; Length 254;  
XX Best Local Similarity 100.0%; Pred. No. 6e-96; Indels 0; Gaps 0;  
XX Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATAGNPMGWLGLYLILGVAGSLVSGSCSIINGEDCSFHSQPMQAAALVMENELFCSGVL 60  
DB 1 MATAGNPMGWLGLYLILGVAGSLVSGSCSIINGEDCSFHSQPMQAAALVMENELFCSGVL 60  
QY 61 VHPQWVLSNAHCFOFNSYITIGLHSLADBPQSGQVSEASLVRHPEYNRPPLANDLMLI 120  
DB 61 VHPQWVLSNAHCFOFNSYITIGLHSLADBPQSGQVSEASLVRHPEYNRPPLANDLMLI 120  
QY 121 KLDESVSSEDTIRISIASQCFTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSSEVCSK 180  
DB 121 KLDESVSSEDTIRISIASQCFTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSSEVCSK 180  
QY 181 LYDPLYPHSMFCAGGQKQKSCNDSGGPLICNGYLGSLVSGFKAPCGQVGVYTNL 240  
DB 181 LYDPLYPHSMFCAGGQKQKSCNDSGGPLICNGYLGSLVSGFKAPCGQVGVYTNL 240  
QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254  
XX  
XX RESULT 11  
XX AAB74830  
XX ID AAB74830 standard; protein; 1079 AA.  
XX  
XX AC AAB74830;  
XX  
XX DT 14-JUN-2001 (first entry)  
XX  
XX DE Prostate tumour antigen amino acid sequence for a fusion protein.  
XX  
XX KM Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;  
XX prostate cancer; immunogenic; cytostatic; vaccine.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200125272-A2.  
XX  
XX PD 12-APR-2001.  
XX  
XX PF 04-OCT-2000; 2000WO-US027464.  
XX  
XX PR 04-OCT-1999; 99US-0157455P.  
XX  
XX PA (CORI-) CORIXA CORP.  
XX  
XX PI Xu J, Skeiky YAW, Reed SG, Cheever MA;  
XX WPI; 2001-245062/25.  
XX  
XX PT Prostate specific protein and its encoding polynucleotide, useful for the  
XX treatment and diagnosis of prostate cancer.  
XX  
XX PS Disclosure; Page 272-276; 276pp; English.  
XX

CC The present invention describes an isolated polypeptide (I) comprising at  
 CC least an immunogenic portion of a prostate tumour antigen protein or its  
 CC variant. (I) have cytostatic activity and can be used in vaccine  
 CC production. (I), prostate tumour antigen polynucleotides, an antigen  
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a  
 CC pharmaceutical composition containing (I) are useful for inhibiting the  
 CC development of cancer in a patient. Antibodies specific for prostate  
 CC specific proteins and oligonucleotides that hybridise to a polynucleotide  
 CC that encodes a prostate specific protein are useful for detecting the  
 CC presence or absence of a cancer or monitoring the progression the  
 CC progression of a cancer, especially prostate cancer. AAH02422 to AAH2872,  
 CC AA874798 to AA874821 and AA874830 are sequences used in the  
 CC exemplification of the present invention

XX Sequence 1079 AA;

Query Match 100.0%; Score 1369; DB 4; Length 1079;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-95;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGNFWGFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 60  
 DB 271 MATAGNFWGFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 330

QY 61 VHPQWVLSAACHFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 120  
 DB 331 VHPQWVLSAACHFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 390

QY 121 KLDSESVESDITRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
 DB 391 KLDSESVESDITRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 450

QY 181 LYDPLYHPSMFCAGGQDQKDCNCGDSCGGLICNGYLQGLVSGFKAPCGQGVGVGYVTNL 240  
 DB 451 LYDPLYHPSMFCAGGQDQKDCNCGDSCGGLICNGYLQGLVSGFKAPCGQGVGVGYVTNL 510

QY 241 CKFTEWIEKTVQAS 254  
 DB 511 CKFTEWIEKTVQAS 524

RESULT 12

ABU71860

ID ABU71860 standard; protein; 1079 AA.

AC ABU71860;

XX 10-JUN-2003 (first entry)

DE Prostate specific antigen fusion protein #2.

KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 KW PSMA.

XX Homo sapiens.

OS Synthetic.

XX US2002192763-A1.

XX 19-DEC-2002.

PF 29-JUN-2001; 2001US-00895793.

XX 04-OCT-1999; 99US-0157455P.

PR 04-OCT-2000; 2000US-00679272.

PR 28-MAR-2001; 2001US-00822827.

XX (XUJ/) XU J.

PA (DILL/) DILLON D C.

PA (MITC/) MITCHAM J L.

PA (HARL/) HARLOCKER S L.

PA (JIAN/) JIANG Y.  
 PA (KALO/) KALOS M D.  
 PA (FANG/) FANGER G R.  
 PA (RETT/) RETTER M W.  
 PA (STOL/) STOLK J A.  
 PA (DAYC/) DAY C H.  
 PA (VEDV/) VEDVICK T S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y A W.  
 PA (HEPL/) HEPLER W T.  
 PA (HEND/) HENDERSON R A.  
 PA (HURA/) HURAL J.  
 PA (MCNE/) MCNEILL P D.  
 PA (HOUG/) HOUGHTON R L.  
 PA (DBAS/) Y DE BASSOLS C V.  
 PA (FOYT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA, Hural J;  
 PI McNeill PD, Houghton RL, Y De BassolsCV, Foy TM;

XX WPI; 2001-245062/25.

XX Prostate specific protein and its encoding polynucleotide, useful for the  
 PT treatment and diagnosis of prostate cancer.

XX Claim 5; SEQ ID NO 947; 85pp; English.

XX The invention describes a fusion protein comprising at least one amino  
 CC acid sequence of immunogenic portions of any of the 3 sequences not  
 CC defined in the specification, or sequences having at least 70 or 90 %  
 CC sequence identity to any one of the 35 sequences defined in the USPTO web  
 CC site, which is encoded by any of the 4 nucleotide sequences not defined  
 CC in the specification. The fusion protein, composition and methods are  
 CC useful for diagnosing, preventing and/or treating cancer, particularly  
 CC prostate cancer. The proteins are useful as markers to indicate the  
 CC presence or absence of cancer. This is the amino acid sequence of a  
 CC fusion protein of the invention created from fragments of prostate  
 CC specific antigen (PSA), prostatic acid phosphatase (PAP), prostatic  
 CC specific membrane antigen (PSMA) and prostate specific antigens of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from the US patent office at  
 CC seqdata.uspto.gov/sequence.html?DocID=US20020192763

XX Sequence 1079 AA;

Query Match 100.0%; Score 1369; DB 4; Length 1079;

Best Local Similarity 100.0%; Pred. No. 2.6e-95;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGNFWGFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 60  
 DB 271 MATAGNFWGFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 330

QY 61 VHPQWVLSAACHFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 120  
 DB 331 VHPQWVLSAACHFQNSYITIGLHSLVAGSLVSGSCQIINGEDCSPHSQPQWQALVMENELFCSGVL 390

QY 121 KLDSESVESDITRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
 DB 391 KLDSESVESDITRISISASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 450

QY 181 LYDPLYHPSMFCAGGQDQKDCNCGDSCGGLICNGYLQGLVSGFKAPCGQGVGVGYVTNL 240  
 DB 451 LYDPLYHPSMFCAGGQDQKDCNCGDSCGGLICNGYLQGLVSGFKAPCGQGVGVGYVTNL 510

QY 241 CKFTEWIEKTVQAS 254  
 DB 511 CKFTEWIEKTVQAS 524



RESULT 13  
AAB21320  
ID AAB21320 standard; protein; 254 AA.  
XX  
AC AAB21320;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human prostate.  
XX  
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; prostate;  
KW kallikrein-like protein; serine protease; cytostatic; cancer;  
KW prostate cancer.  
XX  
OS Homo sapiens.  
XX  
PN W0200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
XX 09-MAR-2000; 2000WO-CA000258.  
XX  
XX 11-MAR-1999; 99US-0124260P.  
XX  
XX 01-APR-1999; 99US-0127386P.  
XX  
XX 21-JUL-1999; 99US-0144919P.  
XX  
XX (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
XX Yousef GM, Diamandis EP;  
XX  
XX WPI; 2000-587440/55.  
XX  
XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
XX protein mediated disorders, especially cancer.  
XX  
XX Example 4; Fig 17; 184pp; English.  
XX  
XX The present sequence is human prostate, a member of the serine protease  
XX family. Kallikreins and kallikrein-like proteins are a subgroup of the  
XX serine protease enzyme family. They catalyse the selective cleavage of  
XX specific polypeptide precursors to release peptides with potent  
XX biological activity. Nucleic acids encoding kallikrein-like proteins KLK-  
XX L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The  
XX proteins are useful in the treatment, monitoring and diagnosis of  
XX cancers, especially prostate cancer. They can also be used to identify a  
XX substance that can associate with or mediate the biological activity of  
XX the proteins. Antibodies can be used to treat conditions mediated by the  
XX kallikrein-like proteins  
XX  
SQ Sequence 254 AA;  
Query Match 99.6%; Score 1364; DB 3; Length 254;  
Best Local Similarity 99.6%; Pred. No. 1.4e-95;  
Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MATAGNPMGFWLGLVILGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60  
DB 1 MATAGNPMGFWLGLVILGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60  
QY 61 VHPQVTL SAHCFQNSVTIIGLHSLBADQPGSQMVEASLSVRHPEYNRPPLANDLMLI 120  
DB 61 VHPQVTL SAHCFQNSVTIIGLHSLBADQPGSQMVEASLSVRHPEYNRPPLANDLMLI 120  
QY 121 KLDSVESSEDTIRISIASQCPAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
DB 121 KLDSVESSEDTIRISIASQCPAGNSCLVSGWGLLANGRMPTVLCQNVSVVSEVCSK 180  
QY 181 LYDPLXHPNFCAGGQDQKDCNNGSGGPLICNGYLQGLVSPGKAPCGQGVPGVYTNL 240  
DB 181 LYDPLXHPNFCAGGQDQKDCNNGSGGPLICNGYLQGLVSPGKAPCGQGVPGVYTNL 240

QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254  
RESULT 14  
AAY72525  
ID AAY72525 standard; protein; 254 AA.  
XX  
AC AAY72525;  
XX  
DT 02-MAY-2001 (first entry)  
XX  
DE Human prostate antigen #3.  
XX  
KW Human; prostate; P703P; gene therapy; prostate cancer; cytostatic; PIN;  
KW prostate intraepithelial neoplasia; NSI; non-structural protein; vaccine;  
KW haemagglutinin; prostatic hyperplasia; prostate-specific serine protease;  
KW therapy.  
XX  
OS Homo sapiens.  
XX  
PN W0200104143-A2.  
XX  
PD 18-JAN-2001.  
XX  
PP 11-JUL-2000; 2000WO-EP006618.  
XX  
XX 13-JUL-1999; 99US-00352616.  
XX  
XX 12-NOV-1999; 99US-00439313.  
XX  
XX 18-NOV-1999; 99US-00443686.  
XX  
XX 14-JAN-2000; 2000US-00483672.  
XX  
XX 27-MAR-2000; 2000US-00536857.  
XX  
XX 09-MAY-2000; 2000US-00568100.  
XX  
XX 12-MAY-2000; 2000US-00570737.  
XX  
XX 13-JUN-2000; 2000US-00593793.  
XX  
XX 27-JUN-2000; 2000GB-00015747.  
XX  
XX 27-JUN-2000; 2000US-00605783.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Cabezon Silva T, Dillon DC;  
XX  
XX WPI; 2001-159380/16.  
XX  
XX New prostate protein or its fragment linked to a immunological/expression  
XX enhancer fusion partner, useful for preparing vaccines for treating  
XX prostate cancers or prostate associated tumors.  
XX  
XX Claim 2; Page 49; 50pp; English.  
XX  
XX The present protein sequence is a human prostate antigen. The invention  
XX relates to a prostate protein (prostate-specific serine protease) or its  
XX fragment fused to an immunological or an expression enhancer fusion  
XX partner, such as NSI (haemagglutinin). The fusion protein is used for  
XX formulating vaccines useful for immunotherapeutically treating patients  
XX susceptible to or suffering from prostate-cancer and prostate-expressing  
XX tumours other than prostate tumours, prostatic hyperplasia and prostate  
XX intraepithelial neoplasia (PIN). Genetic constructs containing prostate  
XX nucleic acids are also useful in gene therapy techniques to induce an  
XX immune response  
XX  
SQ Sequence 254 AA;  
Query Match 99.6%; Score 1364; DB 4; Length 254;  
Best Local Similarity 99.6%; Pred. No. 1.4e-95;  
Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MATAGNPMGFWLGLVILGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60  
DB 1 MATAGNPMGFWLGLVILGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60

QY 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPPLANDMLI 120  
DB 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPPLANDMLI 120  
QY 121 KLDESVSSESDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLCVNVSVVSEEVCSK 180  
DB 121 KLDESVSSESDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLCVNVSVVSEEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKDSNGDSGGPLICNGYLQGLVSGFKAPCGQGVGVVYTNL 240  
DB 181 LYDPLYHPSMFCAGGQDQKDSNGDSGGPLICNGYLQGLVSGFKAPCGQGVGVVYTNL 240  
QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254

## RESULT 15

AAU74901  
ID AAU74901 standard; protein: 254 AA.

AC AAU74901;

XX 09-APR-2002 (first entry)

XX Protein sequence of prostate homologue #3.

XX Human; cytostatic; immunostimulant; vaccine; prostate antigen;  
XX prostate cancer; prostate-specific serine protease;  
XX prostate-associated tumour.

XX Homo sapiens.

XX WO200200708-A2.

XX 03-JAN-2002.

XX 21-JUN-2001; 2001WO-EP007079.

XX 27-JUN-2000; 2000GB-00015736.

XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

XX Cabezón Silva TEV, Marchand M, Vinals Y De BassolsC;

XX WPI; 2002-147879/19.

XX Prostate antigen derivative carrying a mutation in the active site useful  
XX manufacture of a vaccine for treating a patient suffering from prostate  
XX cancer or other prostate-associated tumours.

XX Disclosure; Page 62; 65pp; English.

XX The present invention relates to a new prostate antigen derivative where  
XX the prostate sequence is carrying a mutation in the active site of the  
XX protein. The molecules of the invention may be used for the manufacture  
XX of a vaccine for treating a patient suffering from prostate cancer or  
XX other prostate-associated tumours. The present amino acid sequence  
XX represents one of a collection of prostate homologues (AAU74769-AAU74770  
XX and AAU74901-AAU74902) of the invention. The homologues were used in the  
XX invention to create mutant prostate sequences (AAU74767 and AAU74903)

XX Sequence 254 AA;

Query Match 99.6%; Score 1364; DB 5; Length 254;

Best Local Similarity 99.6%; Pred. No. 1.4e-95;

Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MATAGNPNWFLGYLLVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGYL 60

DB 1 MATAGNPNWFLGYLLVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGYL 60

QY 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPPLANDMLI 120

DB 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPPLANDMLI 120  
QY 121 KLDESVSSESDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLCVNVSVVSEEVCSK 180  
DB 121 KLDESVSSESDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLCVNVSVVSEEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKDSNGDSGGPLICNGYLQGLVSGFKAPCGQGVGVVYTNL 240  
DB 181 LYDPLYHPSMFCAGGQDQKDSNGDSGGPLICNGYLQGLVSGFKAPCGQGVGVVYTNL 240  
QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254

Search completed: May 5, 2004, 16:07:56

Job time : 62 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1369	100.0	254	9	US-09-759-143-525	Sequence 525, App
2	1369	100.0	254	9	US-09-780-663-525	Sequence 525, App
3	1369	100.0	254	9	US-09-822-827-525	Sequence 525, App
4	1369	100.0	254	9	US-09-899-793-525	Sequence 525, App
5	1369	100.0	254	9	US-09-899-874-525	Sequence 525, App
6	1369	100.0	254	13	US-10-012-896-525	Sequence 525, App
7	1369	100.0	254	14	US-10-101-940-525	Sequence 525, App
8	1369	100.0	254	14	US-10-144-678A-525	Sequence 525, App
9	1369	100.0	254	14	US-10-1294-025-525	Sequence 525, App
10	1369	100.0	1079	9	US-09-822-827-947	Sequence 947, App
11	1369	100.0	1079	9	US-09-899-793-947	Sequence 947, App
12	1364	99.6	254	14	US-10-312-089-947	Sequence 947, App
13	1352	98.8	254	9	US-09-759-143-523	Sequence 523, App
14	1352	98.8	254	9	US-09-780-663-523	Sequence 523, App
15	1352	98.8	254	9	US-09-822-827-523	Sequence 523, App

QY 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
Db 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
Db 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
QY 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
Db 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
Db 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
QY 241 CKFTEWIEKTVQAS 254  
Db 241 CKFTEWIEKTVQAS 254

## RESULT 2

US-09-780-669-525

; Sequence 525, Application US/09780669

; Patent No. US20020051977A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqui

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John H.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.427C24

; CURRENT FILING DATE: US/09/780,669

; NUMBER OF SEQ ID NOS: 943

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 525

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-780-669-525

Query Match 100.0%; Score 1369; DB 9; Length 254;  
Best Local Similarity 100.0%; Pred. No. 4.7e-131;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
Db 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
Db 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
QY 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Db 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
Db 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
QY 241 CKFTEWIEKTVQAS 254  
Db 241 CKFTEWIEKTVQAS 254

## RESULT 3

US-09-822-827-525

; Sequence 525, Application US/09822827

; Patent No. US20020081680A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.534C1

; CURRENT FILING DATE: US/09/822,827

; NUMBER OF SEQ ID NOS: 982

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 525

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Homo sapien

US-09-822-827-525

Query Match 100.0%; Score 1369; DB 9; Length 254;  
Best Local Similarity 100.0%; Pred. No. 4.7e-131;  
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
Db 1 MATAGNPNWFLGYLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
Db 61 VHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGSGQWVEASLSVRHPEYNRPPLANDLMLI 120  
QY 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
Db 121 KLDSESVESDTRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
Db 181 LYDPLYHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
QY 241 CKFTEWIEKTVQAS 254  
Db 241 CKFTEWIEKTVQAS 254

## RESULT 4

US-09-895-793-525

; Sequence 525, Application US/09895793

; Publication No. US20020192763A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqui

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John H.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-793-525

Query Match 100.0%; Score 1369; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGPMWFLGYLLIGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60
DB 1 MATAGPMWFLGYLLIGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGQSQMVEASLSVRHPEYNRPILLANDLMLI 120
DB 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGQSQMVEASLSVRHPEYNRPILLANDLMLI 120
QY 121 KLDSESVESDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEVCSK 180
DB 121 KLDSESVESDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEVCSK 180
QY 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVGYTNL 240
DB 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVGYTNL 240
QY 241 CKTEWIEKTVQAS 254
DB 241 CKTEWIEKTVQAS 254

RESULT 5
US-09-895-814-525
; Sequence 525, Application US/09895814
; Publication No. US2002013296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meascher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896

; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-895-814-525

Query Match 100.0%; Score 1369; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATAGPMWFLGYLLIGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60
DB 1 MATAGPMWFLGYLLIGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGVL 60
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGQSQMVEASLSVRHPEYNRPILLANDLMLI 120
DB 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGQSQMVEASLSVRHPEYNRPILLANDLMLI 120
QY 121 KLDSESVESDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEVCSK 180
DB 121 KLDSESVESDTIRISIASQCTAGNSCLVSGWGLLANGRMPTVLCQVNVSVVSEVCSK 180
QY 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVGYTNL 240
DB 181 LYDPLVHPSMFCAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVGYTNL 240
QY 241 CKTEWIEKTVQAS 254
DB 241 CKTEWIEKTVQAS 254

RESULT 6
US-10-012-896-525
; Sequence 525, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meascher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012.896
```

```
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-896-525

Query Match      100.0%; Score 1369; DB 13; Length 254;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60
Db 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60

Qy 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120
Db 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120

Qy 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240
Db 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240

Qy 241 CKFTWIEKTVQAS 254
Db 241 CKFTWIEKTVQAS 254

RESULT 7
US-10-010-940-525
; Sequence 525, Application US/10010940
; Publication No. US20030089062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqun
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-010-940-525

Query Match      100.0%; Score 1369; DB 14; Length 254;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60
Db 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60

Qy 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120
Db 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120

Qy 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240
Db 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240

Qy 241 CKFTWIEKTVQAS 254
Db 241 CKFTWIEKTVQAS 254

RESULT 8
US-10-144-678A-525
; Sequence 525, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John H.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yashir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Bassols, Carlota
; APPLICANT: Foy, Teresa M.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C28
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-525

Query Match      100.0%; Score 1369; DB 14; Length 254;
Best Local Similarity 100.0%; Pred. No. 4.7e-131;
Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60
Db 1 MATAGNPGWFLGYLLGVAGSLVSGSCQIINGEDCSPHSQPQWQAALVMENELFCSGYL 60

Qy 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120
Db 61 VHPQWLSAAHCFQNSYITIGLHLSLEADQPGSQMVEASLSVRHPEYNRPPLANDMLI 120

Qy 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDSEVSSEDTRISISIAQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240
Db 181 LYDPLXHPSMFCAGGQDQKDCNGDSGGPLICNGYLQGLVSGKAPCGQGVPGVYTNL 240

Qy 241 CKFTWIEKTVQAS 254
Db 241 CKFTWIEKTVQAS 254
```



QY 61 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 120  
 Db 331 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 330  
 QY 121 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180  
 Db 391 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 450  
 QY 181 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
 Db 451 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 510  
 QY 241 CKFTEWIEKTVQAS 254  
 Db 511 CKFTEWIEKTVQAS 524

## RESULT 12

US-10-312-089-7  
 ; Sequence 7, Application US/10312089  
 ; Publication No. US20030143240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Caberon-Silva, Teresa Elisa Virginia  
 ; APPLICANT: Permann, Philippe Jean Gervais Ghislain  
 ; TITLE OF INVENTION: Prostate Protein Vaccine Comprising  
 ; TITLE OF INVENTION: Derivatised Thiol Residues and Methods for Producing Said  
 ; TITLE OF INVENTION: Antigen  
 ; FILE REFERENCE: B45224  
 ; CURRENT APPLICATION NUMBER: US/10/312,089  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: PCT/E01/07082  
 ; PRIOR FILING DATE: 2001-06-21  
 ; PRIOR APPLICATION NUMBER: GB 0015722.2  
 ; PRIOR FILING DATE: 2000-06-27  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 254  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 ; US-10-312-089-7

Query Match 99.6%; Score 1364; DB 14; Length 254;  
 Best Local Similarity 99.6%; Pred. No. 1.5e-130;  
 Matches 253; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MATAGNPMGWFLGYLILGVAGSLVSGSCQIINGDCSPHSPQWQAALVMENELFCSGVL 60  
 Db 1 MATAGNPMGWFLGYLILGVAGSLVSGSCQIINGDCSPHSPQWQAALVMENELFCSGVL 60  
 QY 61 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 120  
 Db 61 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 120  
 QY 121 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180  
 Db 121 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180  
 QY 181 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
 Db 181 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
 QY 241 CKFTEWIEKTVQAS 254  
 Db 241 CKFTEWIEKTVQAS 524

## RESULT 13

US-09-759-143-523  
 ; Sequence 523, Application US/09759143  
 ; Patent No. US2002022248A1  
 ; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Mitcham, Jennifer L.  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Jiang, Yuqi  
 APPLICANT: Henderson, Robert A.  
 APPLICANT: Kalos, Michael D.  
 APPLICANT: Fanger, Gary R.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Stolk, John A.  
 APPLICANT: Day, Craig H.  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Carter, Darrick  
 APPLICANT: Li, Samuel  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Hepler, William  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 FILE REFERENCE: 210121.427C23  
 CURRENT APPLICATION NUMBER: US/09/759,143  
 CURRENT FILING DATE: 2001-01-12  
 NUMBER OF SEQ ID NOS: 934  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 523  
 LENGTH: 254  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Made in a lab  
 NAME/KEY: VARIANT  
 LOCATION: (1)...(254)  
 OTHER INFORMATION: Xaa = any amino acid  
 US-09-759-143-523

Query Match 98.8%; Score 1352; DB 9; Length 254;  
 Best Local Similarity 98.8%; Pred. No. 2.6e-129;  
 Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MATAGNPMGWFLGYLILGVAGSLVSGSCQIINGDCSPHSPQWQAALVMENELFCSGVL 60  
 Db 1 MATAGNPMGWFLGYLILGVAGSLVSGSCQIINGDCSPHSPQWQAALVMENELFCSGVL 60  
 QY 61 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 120  
 Db 61 VHPQWLSAACHFQNSYTTIGLHLSLEADQEPGSMVEASLSVRHPEYNRPLLANDLMLI 120  
 QY 121 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180  
 Db 121 KLDSESVESDTRISISASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180  
 QY 181 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
 Db 181 LYDPLVHPMFAGGQDQKSCNGDSGGPLICNGYLQGLVSGKAPCGQVGPVYTNL 240  
 QY 241 CKFTEWIEKTVQAS 254  
 Db 241 CKFTEWIEKTVQAS 254

## RESULT 14

US-09-780-669-523  
 ; Sequence 523, Application US/09780669  
 ; Patent No. US20020051977A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yuqi  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.



APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Repler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780.669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 523  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Made in a lab  
NAME/KEY: VARIANT  
LOCATION: (1)...(254)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-780-669-523

Query Match 98.8%; Score 1352; DB 9; Length 254;  
Best Local Similarity 98.8%; Pred. No. 2.6e-129;  
Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATAGNPMGWFLGYLILGVAGSLVSGSCSIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
DB 1 MATAGNPMGWFLGYLILGVAGSLVSGSCSIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPLANDMLI 120  
DB 61 VHPQWVLSATHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPLANDMLI 120  
QY 121 KLDSESVESDITRSISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
DB 121 KLDSESVESDITRSISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGGQDKSCNGDSGGPLICNGYLOGLVSPGKAPCGQVGPVYTNL 240  
DB 181 LYDPLYHPSMFCAGGGQDKSCNGDSGGPLICNGYLOGLVSPGKAPCGQVGPVYTNL 240  
QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254

Search completed: May 5, 2004, 16:15:19  
Job time : 48 secs

APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Repler, William  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C24  
CURRENT APPLICATION NUMBER: US/09/780.669  
CURRENT FILING DATE: 2001-02-09  
NUMBER OF SEQ ID NOS: 943  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 523  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Made in a lab  
NAME/KEY: VARIANT  
LOCATION: (1)...(254)  
OTHER INFORMATION: Xaa = any amino acid  
US-09-780-669-523

Query Match 98.8%; Score 1352; DB 9; Length 254;  
Best Local Similarity 98.8%; Pred. No. 2.6e-129;  
Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATAGNPMGWFLGYLILGVAGSLVSGSCSIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
DB 1 MATAGNPMGWFLGYLILGVAGSLVSGSCSIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAAHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPLANDMLI 120  
DB 61 VHPQWVLSATHCFQNSYITIGLHLSLEADQEPGSMQVEASLSVRHPEYNRPLANDMLI 120  
QY 121 KLDSESVESDITRSISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
DB 121 KLDSESVESDITRSISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180  
QY 181 LYDPLYHPSMFCAGGGQDKSCNGDSGGPLICNGYLOGLVSPGKAPCGQVGPVYTNL 240  
DB 181 LYDPLYHPSMFCAGGGQDKSCNGDSGGPLICNGYLOGLVSPGKAPCGQVGPVYTNL 240  
QY 241 CKFTIEWIEKTVQAS 254  
DB 241 CKFTIEWIEKTVQAS 254

Search completed: May 5, 2004, 16:15:19  
Job time : 48 secs

RESULT 15  
US-09-822-827-523  
Sequence 523, Application US/09822827  
Patent No. US20020081680A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.534C1  
CURRENT APPLICATION NUMBER: US/09/822.827  
CURRENT FILING DATE: 2001-03-28  
NUMBER OF SEQ ID NOS: 982  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 523  
LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 16:06:47 ; Search time 23 Seconds  
(without alignments)  
570.130 Million cell updates/sec

Title: US-09-895-814-525

Perfect score: 1369

Sequence: 1 MATAGNPGWFLGILGVA.....GVYTNLCKFTWERTVQAS 254

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1369	100.0	254	4	US-09-439-313-525
2	1369	100.0	254	4	US-09-636-215-525
3	1369	100.0	254	4	US-09-685-166A-525
4	1352	98.8	254	4	US-09-439-313-523
5	1352	98.8	254	4	US-09-636-215-523
6	1352	98.8	254	4	US-09-685-166A-523
7	1205	88.0	449	4	US-09-636-215-617
8	1205	88.0	449	4	US-09-685-166A-617
9	1185	86.6	220	4	US-09-439-313-327
10	1185	86.6	220	4	US-09-352-616A-327
11	1185	86.6	220	4	US-09-233-149A-327
12	1185	86.6	220	4	US-09-636-215-327
13	1185	86.6	220	4	US-09-685-166A-327
14	1185	86.6	220	4	US-09-688-489-327
15	1167.5	85.3	284	4	US-09-386-642-54
16	1160	84.7	248	1	US-08-744-026-1
17	1160	84.7	248	1	US-09-102-732-1
18	1160	84.7	248	3	US-09-261-767-1
19	1124	82.1	232	2	US-08-897-340-31
20	1124	82.1	232	3	US-09-252-329-31
21	1070	78.2	205	3	US-09-020-956-176
22	1070	78.2	205	3	US-09-030-607-176
23	1070	78.2	205	4	US-09-439-313-176
24	1070	78.2	205	4	US-09-352-616A-176
25	1070	78.2	205	4	US-09-233-149A-176
26	1070	78.2	205	4	US-09-159-812-176
27	1070	78.2	205	4	US-09-636-215-176

28 1070 78.2 205 4 US-09-685-166A-176 Sequence 176, App  
29 1070 78.2 205 4 US-09-115-453-176 Sequence 176, App  
30 1070 78.2 205 4 US-09-688-489-176 Sequence 176, App  
31 835 61.0 159 3 US-09-020-956-172 Sequence 172, App  
32 835 61.0 159 3 US-09-030-607-172 Sequence 172, App  
33 835 61.0 159 4 US-09-439-313-172 Sequence 172, App  
34 835 61.0 159 4 US-09-352-616A-172 Sequence 172, App  
35 835 61.0 159 4 US-09-233-149A-172 Sequence 172, App  
36 835 61.0 159 4 US-09-159-812-172 Sequence 172, App  
37 835 61.0 159 4 US-09-636-215-172 Sequence 172, App  
38 835 61.0 159 4 US-09-685-166A-172 Sequence 172, App  
39 835 61.0 159 4 US-09-115-453-172 Sequence 172, App  
40 835 61.0 159 4 US-09-688-489-172 Sequence 838, App  
41 830 60.6 166 4 US-09-636-215-838 Sequence 838, App  
42 830 60.6 166 4 US-09-685-166A-838 Sequence 2, Appli  
43 671 49.0 293 4 US-09-509-908-2 Sequence 5522, Ap  
44 584.5 42.7 118 4 US-09-621-976-5522 Sequence 1, Appli  
45 581 42.4 268 2 US-08-824-874-1

#### ALIGNMENTS

RESULT 1  
US-09-439-313-525  
; Sequence 525, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqi  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark  
; APPLICANT: Solk, John  
; APPLICANT: Day, Craig  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C9  
; CURRENT APPLICATION NUMBER: US/09/439,313  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 575  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 525  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-439-313-525

Query Match 100.0%; Score 1369; DB 4; Length 254;  
Best Local Similarity 100.0%; Pred. No. 2.7e-132; Indels 0; Gaps 0;  
Matches 254; Conservative 0; Mismatches 0;

QY 1 MATAGNPGWFLGILGVLGAGSLVSGSCQIINGDCSPHSQFQWQAALVMENELFCSGYL 60  
Db 1 MATAGNPGWFLGILGVLGAGSLVSGSCQIINGDCSPHSQFQWQAALVMENELFCSGYL 60  
QY 61 VHPQWLSAAHCFQNSYTTIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDMLI 120  
Db 61 VHPQWLSAAHCFQNSYTTIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDMLI 120  
QY 121 KLDSESVESDITIRISISIASQCFTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
Db 121 KLDSESVESDITIRISISIASQCFTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
QY 181 LYDPLVHPSMFCAGGGQDKSCNGDSGGPLICNGVLQGLVSGKAPCGQGVGYTNL 240  
Db 181 LYDPLVHPSMFCAGGGQDKSCNGDSGGPLICNGVLQGLVSGKAPCGQGVGYTNL 240  
QY 241 CKFTEWIENTVQAS 254

```

; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-685-166A-525

Query Match 100.0%; Score 1369; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.7e-132; Indels 0; Gaps 0;
Matches 254; Conservative 0; Mismatches 0;

Qy 1 MATAGNPGWFLGYLILGVAGSLVSGSCSQIINGEDCSPHSQPWOQAALWMENELFCSGVL 60
Db 1 MATAGNPGWFLGYLILGVAGSLVSGSCSQIINGEDCSPHSQPWOQAALWMENELFCSGVL 60

Qy 61 VHPQWVLSAAHCFQNSYITIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDLMLI 120
Db 61 VHPQWVLSAAHCFQNSYITIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDLMLI 120

Qy 121 KLDESVSSEDTIRSIASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDESVSSEDTIRSIASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 181 LYDPLYHPSMFAGGGQDQKDSNGDSGGPLICNGYLGSLVFGKAPCGQGVGVYTNL 240
Db 181 LYDPLYHPSMFAGGGQDQKDSNGDSGGPLICNGYLGSLVFGKAPCGQGVGVYTNL 240

Qy 241 CKFTEWIEKTVOAS 254
Db 241 CKFTEWIEKTVOAS 254

RESULT 4
US-09-439-313-523
; Sequence 523, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575

```

```

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-636-215-525

Query Match 100.0%; Score 1369; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.7e-132; Indels 0; Gaps 0;
Matches 254; Conservative 0; Mismatches 0;

Qy 1 MATAGNPGWFLGYLILGVAGSLVSGSCSQIINGEDCSPHSQPWOQAALWMENELFCSGVL 60
Db 1 MATAGNPGWFLGYLILGVAGSLVSGSCSQIINGEDCSPHSQPWOQAALWMENELFCSGVL 60

Qy 61 VHPQWVLSAAHCFQNSYITIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDLMLI 120
Db 61 VHPQWVLSAAHCFQNSYITIGLHSLADQEPGQWVEASLSVRHPEYNRPILLANDLMLI 120

Qy 121 KLDESVSSEDTIRSIASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180
Db 121 KLDESVSSEDTIRSIASQCTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEVCSK 180

Qy 181 LYDPLYHPSMFAGGGQDQKDSNGDSGGPLICNGYLGSLVFGKAPCGQGVGVYTNL 240
Db 181 LYDPLYHPSMFAGGGQDQKDSNGDSGGPLICNGYLGSLVFGKAPCGQGVGVYTNL 240

Qy 241 CKFTEWIEKTVOAS 254
Db 241 CKFTEWIEKTVOAS 254

RESULT 3
US-09-685-166A-525
; Sequence 525, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.

```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-439-313-523

Query Match          98.8%; Score 1352; DB 4; Length 254;
Best Local Similarity 98.8%; Pred. No. 1.5e-130;
Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATAGNPGWFLGYLLILGVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGVL 60
DB 1 MATAGNPGWFLGYLLILGVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGVL 60

QY 61 VHPQWLAAHCFQNSYTIIGLHSLADQEPGSMQVEASLSVRHPEYNRPPLANDLMLI 120
DB 61 VHPQWLAAHCFQNSYTIIGLHSLADQEPGSMQVEASLSVRHPEYNRPPLANDLMLI 120

QY 121 KLDSEVSSEDTIRISIASOCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180
DB 121 KLDSEVSSEDTIRISIASOCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180

QY 181 LYDPLHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNL 240
DB 181 LYDPLHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNL 240

QY 241 CKFTEWIEKTVQAS 254
DB 241 CKFTEWIEKTVQAS 254

RESULT 5
US-09-636-215-523
; Sequence 523, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-685-166A-523

Query Match          98.8%; Score 1352; DB 4; Length 254;
Best Local Similarity 98.8%; Pred. No. 1.5e-130;
Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATAGNPGWFLGYLLILGVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGVL 60
DB 1 MATAGNPGWFLGYLLILGVAGSLVSGSCSIINGEDCSPHSQPWQAALVMENELFCSGVL 60

QY 61 VHPQWLAAHCFQNSYTIIGLHSLADQEPGSMQVEASLSVRHPEYNRPPLANDLMLI 120
DB 61 VHPQWLAAHCFQNSYTIIGLHSLADQEPGSMQVEASLSVRHPEYNRPPLANDLMLI 120

QY 121 KLDSEVSSEDTIRISIASOCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180
DB 121 KLDSEVSSEDTIRISIASOCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVVSEEVCSK 180

QY 181 LYDPLHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNL 240
DB 181 LYDPLHPSMFCAGGQDQKDCNCGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNL 240

QY 241 CKFTEWIEKTVQAS 254
DB 241 CKFTEWIEKTVQAS 254

RESULT 6
US-09-685-166A-523
; Sequence 523, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-685-166A-523

Query Match          98.8%; Score 1352; DB 4; Length 254;
Best Local Similarity 98.8%; Pred. No. 1.5e-130;
Matches 251; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 NATAGNPGWFLGYLLGVAGSLVSGSCSIIINGEDCSFHSQPWQAALVMENELFCSGVL 60  
Db 1 NATAGNPGWFLGYLLGVAGSLVSGSCSIIINGEDCSFHSQPWQAALVMENELFCSGVL 60  
QY 61 VHPQWVLSAACHFQNSYTIIGLHLEADQPPGSMVEASLSVRHPEYRPLLANDMLI 120  
Db 61 VHPQWVLSAACHFQNSYTIIGLHLEADQPPGSMVEASLSVRHPEYRPLLANDMLI 120  
QY 121 KLDESVSSEDTRISISIAOQPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVEEVCVK 180  
Db 121 KLDESVSSEDTRISISIAOQPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVEEVCVK 180  
QY 181 LVDPVLYHPSMFCAGGQDQKSCNDSGGPLICNGYLGSLVFGKAPCGQGVGVVTNL 240  
Db 181 LVDPVLYHPSMFCAGGQDQKSCNDSGGPLICNGYLGSLVFGKAPCGQGVGVVTNL 240  
QY 241 CKFTEWIEKTVOAS 254  
Db 241 CKFTEWIEKTVOAS 254

RESULT 7  
US-09-636-215-617  
; Sequence 617, Application US/09636215  
; Patent No. 6620922  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; APPLICANT: Repl, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.42717C17  
; CURRENT APPLICATION NUMBER: US/09/636,215  
; CURRENT FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 852  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 617  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-636-215-617

Query Match 88.0%; Score 1205; DB 4; Length 449;  
Best Local Similarity 100.0%; Pred. No. 4e-115;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 31 IINGEDCSPHSQPWQAALVMENELFCSGVLVHPQWVLSAACHFQNSYTIIGLHLEADQ 90  
Db 8 IINGEDCSPHSQPWQAALVMENELFCSGVLVHPQWVLSAACHFQNSYTIIGLHLEADQ 67  
QY 91 EPGSQMVEASLSVRHPEYRPLLANDMLIKLDESVSSEDTRISISIAOQPTAGNSCLV 150  
Db 68 EPGSQMVEASLSVRHPEYRPLLANDMLIKLDESVSSEDTRISISIAOQPTAGNSCLV 127  
QY 151 SGWGLLANGRMPTVLQCVNVSVEEVCVKLYDPLVHPSMFCAGGQDQKSCNDSGGP 210  
Db 128 SGWGLLANGRMPTVLQCVNVSVEEVCVKLYDPLVHPSMFCAGGQDQKSCNDSGGP 187  
QY 211 LICNGYLQGLVSGKAPCGQGVGVVTNLCKFTEWIEKTVOAS 254

Db 188 LICNGYLQGLVSGKAPCGQGVGVVTNLCKFTEWIEKTVOAS 231  
RESULT 8  
US-09-685-166A-617  
; Sequence 617, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C21  
; CURRENT APPLICATION NUMBER: US/09/685,166A  
; CURRENT FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 898  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 617  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-685-166A-617

Query Match 88.0%; Score 1205; DB 4; Length 449;  
Best Local Similarity 100.0%; Pred. No. 4e-115;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 31 IINGEDCSPHSQPWQAALVMENELFCSGVLVHPQWVLSAACHFQNSYTIIGLHLEADQ 90  
Db 8 IINGEDCSPHSQPWQAALVMENELFCSGVLVHPQWVLSAACHFQNSYTIIGLHLEADQ 67  
QY 91 EPGSQMVEASLSVRHPEYRPLLANDMLIKLDESVSSEDTRISISIAOQPTAGNSCLV 150  
Db 68 EPGSQMVEASLSVRHPEYRPLLANDMLIKLDESVSSEDTRISISIAOQPTAGNSCLV 127  
QY 151 SGWGLLANGRMPTVLQCVNVSVEEVCVKLYDPLVHPSMFCAGGQDQKSCNDSGGP 210  
Db 128 SGWGLLANGRMPTVLQCVNVSVEEVCVKLYDPLVHPSMFCAGGQDQKSCNDSGGP 187  
QY 211 LICNGYLQGLVSGKAPCGQGVGVVTNLCKFTEWIEKTVOAS 254  
Db 188 LICNGYLQGLVSGKAPCGQGVGVVTNLCKFTEWIEKTVOAS 231

RESULT 9  
US-09-439-313-327  
; Sequence 327, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqi  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary

```

; APPLICANT: Retter, Mark
; APPLICANT: Soik, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439.313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 327
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-439-313-327

Query Match      86.6%; Score 1185; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.6e-113;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 94
DB 1 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 60
QY 95 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 154
DB 61 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 120
QY 155 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 214
DB 121 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 180
QY 215 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
DB 181 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 220

RESULT 10
US-09-352-616A-327
; Sequence 327, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqul
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352.616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 327
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-352-616A-327

Query Match      86.6%; Score 1185; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.6e-113;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 94
DB 1 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 60
QY 95 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 154
DB 61 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 120
QY 155 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 214
DB 121 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 180
QY 215 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
DB 181 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 220

RESULT 11
US-09-232-149A-327
; Sequence 327, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232.149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 327
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-232-149A-327

Query Match      86.6%; Score 1185; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.6e-113;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 94
DB 1 EDCSPHSQPWQAALVMENELFCGVLVHPQWVLSAAHCFQNSYTTIGLHLSLEADQEPGS 60
QY 95 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 154
DB 61 QMVEASLSVRHPEYNRPPLANDMLIKLDESSESDDTIRSIASQCPTAGNSCLVSGWG 120
QY 155 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 214
DB 121 LLANGRMPTVLCQNVSVSVSEVCSKLYDPLVHPMSMFCAGGQDQKDCSNGSGGPLICN 180
QY 215 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
DB 181 GYLQGLVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 220

RESULT 12
US-09-636-215-327
; Sequence 327, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqul
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
```

;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
;; FILE REFERENCE: 210121.42717C17  
;; CURRENT APPLICATION NUMBER: US/09/636,215  
;; CURRENT FILING DATE: 2000-08-10  
;; NUMBER OF SEQ ID NOS: 852  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 327  
;; LENGTH: 220  
;; TYPE: PR1  
;; ORGANISM: Homo sapien  
US-09-636-215-327

Query Match 86.6%; Score 1185; DB 4; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.6e-113;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 94  
Db 1 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 60  
QY 95 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 154  
Db 61 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 120  
QY 155 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 214  
Db 121 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 180  
QY 215 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 254  
Db 181 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 220

## RESULT 13

US-09-685-166A-327  
; Sequence 327, Application US/09685166A  
; Patent No. 6630305

;; GENERAL INFORMATION:  
;; APPLICANT: Xu, Jiangchun

;; APPLICANT: Dillon, Davin C.

;; APPLICANT: Mitcham, Jennifer L.

;; APPLICANT: Harlocker, Susan L.

;; APPLICANT: Jiang, Yuqi

;; APPLICANT: Henderson, Robert A.

;; APPLICANT: Kalos, Michael D.

;; APPLICANT: Fanger, Gary R.

;; APPLICANT: Retter, Marc W.

;; APPLICANT: Stolk, John A.

;; APPLICANT: Day, Craig H.

;; APPLICANT: Vedwick, Thomas S.

;; APPLICANT: Carter, Darrick

;; APPLICANT: Li, Samuel

;; APPLICANT: Wang, Aijun

;; APPLICANT: Skeiky, Yasir A.W.

;; APPLICANT: Hepler, William

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

;; FILE REFERENCE: 210121.427C21

;; CURRENT APPLICATION NUMBER: US/09/685,166A

;; CURRENT FILING DATE: 2000-10-10

;; NUMBER OF SEQ ID NOS: 898

;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 327

;; LENGTH: 220

;; TYPE: PR1

;; ORGANISM: Homo sapien

US-09-685-166A-327

Query Match 86.6%; Score 1185; DB 4; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.6e-113;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 94

Db 1 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 60  
QY 95 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 154  
Db 61 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 120  
QY 155 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 214  
Db 121 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 180  
QY 215 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 254  
Db 181 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 220

## RESULT 14

US-09-688-489-327

;; Sequence 327, Application US/09688489

;; Patent No. 6664377

;; GENERAL INFORMATION:

;; APPLICANT: Xu, Jiangchun

;; APPLICANT: Dillon, Davin C.

;; APPLICANT: Mitcham, Jennifer Lynn

;; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE

;; TITLE OF INVENTION: CANCER AND METHODS FOR THEIR USE

;; FILE REFERENCE: 210121.427D2

;; CURRENT APPLICATION NUMBER: US/09/688,489

;; CURRENT FILING DATE: 2000-10-13

;; NUMBER OF SEQ ID NOS: 338

;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 327

;; LENGTH: 220

;; TYPE: PR1

;; ORGANISM: Homo sapien

US-09-688-489-327

Query Match 86.8%; Score 1185; DB 4; Length 220;  
Best Local Similarity 100.0%; Pred. No. 1.6e-113;  
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 94

Db 1 EDCSPHSQWQAALVWENELFCGVLVHPQWLSAAHCFQNSYTIIGLHLSLEADQEPGS 60

QY 95 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 154

Db 61 QMVEASLSVRHPEYRNPRLANDMLIKLDESVSSEDITRSISASQCTAGNSCLVSGWG 120

QY 155 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 214

Db 121 LLANGRMPTVLCQVNVSVSEVCSKLYDPLVHPMSFCAGGGQDQKDCNCGDSGGPLCN 180

QY 215 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 254

Db 181 GYLOGLVSFGKAPCGQGVPGVYTNLCCKFTEWIEKTVQAS 220

## RESULT 15

US-09-386-642-54

;; Sequence 54, Application US/09386642

;; Patent No. 6420157

;; GENERAL INFORMATION:

;; APPLICANT: Darrow, Andrew

;; APPLICANT: Qi, Jensen

;; APPLICANT: Andrade-Gordon, Patricia

;; TITLE OF INVENTION: Zymogen Activation System

;; FILE REFERENCE: ORT-1028

;; CURRENT APPLICATION NUMBER: US/09/386,642

;; CURRENT FILING DATE: 1999-08-31

;; NUMBER OF SEQ ID NOS: 60

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 54

```
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human MH2
; OTHER INFORMATION: protease in PFEK zymogen vector
US-09-386-642-54

Query Match      85.3%; Score 1167.5; DB 4; Length 284;
Best Local Similarity 96.5%; Pred. No. 1.5e-111;
Matches 218; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 30 QIINGEDC-SPHSQFWQAALVWENELFCSGVLVHPQWVLSAAHCFQNSYTIIGLHLSLEA 88
Db 51 KIVGGYNCLPHSQFWQAALVWENELFCSGVLVHPQWVLSAAHCFQNSYTIIGLHLSLEA 110
QY 89 DQEPGSQWVEASLSVRHPEYNRPILLANDMLIKLDESVSSEDTIRISISIASQCPTAGNSC 148
Db 111 DQEPGSQWVEASLSVRHPEYNRPILLANDMLIKLDESVSSEDTIRISISIASQCPTAGNSC 170
QY 149 LVSGMGLLANGRMPTVLQCVNVSVVSEVCSKLYDPLYHPSPMFCAGGQDQKSCNGDSG 208
Db 171 LVSGMGLLANGRMPTVLQCVNVSVVSEVCSKLYDPLYHPSPMFCAGGQDQKSCNGDSG 230
QY 209 GPLICNGYLGIVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254
Db 231 GPLICNGYLGIVSFGKAPCGQGVPGVYTNLCKFTWIEKTVQAS 276
```

Search completed: May 5, 2004, 16:10:31  
Job time : 24 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 16:04:47 ; Search time 20 seconds  
(without alignments)  
1221.632 Million cell updates/sec

Title: US-09-895-814-525

Perfect score: 1369

Sequence: 1 MATAGNFWGFLGYLILGVA.....GVYTNLCRTEWIKTVQAS 254

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_78.\*

2: PIR1.\*

3: PIR2.\*

4: PIR3.\*

5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567.5	41.5	253	2 A53968	serine proteinase
2	563.5	41.2	260	2 I56559	neuropsin - mouse
3	524	38.3	261	2 A45303	tissue kallikrein
4	515	37.6	244	2 A44284	tissue kallikrein
5	510.5	37.3	248	2 S55066	trypsin (EC 3.4.21)
6	509	37.2	261	1 NGMSG	7S nerve growth fa
7	507	37.0	261	2 A34079	tissue kallikrein
8	506	37.0	247	2 A27547	trypsin (EC 3.4.21)
9	506	37.0	261	2 S01971	tissue kallikrein
10	503	36.7	261	2 J02036	tissue kallikrein
11	503	36.7	261	2 A31136	tissue kallikrein
12	502.5	36.7	246	2 J01472	trypsin (EC 3.4.21)
13	502.5	36.7	261	2 A29586	tissue kallikrein
14	502	36.7	261	2 A41020	tissue kallikrein
15	501.5	36.6	250	2 T01779	trypsin (EC 3.4.21)
16	499.5	36.5	246	1 TRDGC	trypsin (EC 3.4.21)
17	498	36.4	261	1 A32297	semenogelase (EC 3
18	497	36.3	247	2 S05494	trypsin (EC 3.4.21)
19	497	36.3	259	2 B31136	tissue kallikrein
20	496	36.2	248	2 S55067	trypsin (EC 3.4.21)
21	496	36.2	259	2 A29746	tissue kallikrein
22	496	36.2	265	1 K0RTP	tissue kallikrein
23	495.5	36.2	246	2 J01471	trypsin (EC 3.4.21)
24	492.5	36.0	261	2 A29745	tissue kallikrein
25	492	35.9	257	2 S33772	tissue kallikrein
26	491	35.9	261	1 EGMSB	tissue kallikrein
27	488.5	35.7	261	1 A25606	tissue kallikrein
28	487.5	35.6	262	1 K0HU	tissue kallikrein
29	485	35.4	246	1 TRRT1	trypsin (EC 3.4.21)

30	485	35.4	261	1 S35711	semenogelase (EC 3
31	484	35.4	259	1 K0RTTN	tonin (EC 3.4.21.-
32	484	35.4	261	1 TRMSMS	tissue kallikrein
33	480	35.1	247	2 S13813	trypsin (EC 3.4.21
34	480	35.1	261	1 K0MS1	tissue kallikrein
35	479.5	35.0	229	1 TRBOTR	trypsin (EC 3.4.21
36	479	35.0	246	2 B25528	trypsin (EC 3.4.21
37	476.5	34.8	242	2 S49489	trypsin (EC 3.4.21
38	472.5	34.5	231	1 TRPGTR	trypsin (EC 3.4.21
39	472.5	34.5	247	1 TRDG	trypsin (EC 3.4.21
40	471	34.4	259	2 D23863	tissue kallikrein
41	470.5	34.4	263	2 S15686	tissue kallikrein
42	470	34.3	243	2 A35871	tissue kallikrein
43	469.5	34.3	260	2 A37938	tissue kallikrein
44	467.5	34.1	238	2 S31779	trypsin (EC 3.4.21
45	467.5	34.1	247	1 A25852	trypsin (EC 3.4.21

#### ALIGNMENTS

##### RESULT 1

A53968

serine proteinase SCCE precursor - human

N/Alternate names: stratum corneum chymotryptic enzyme

C/Species: Homo sapiens (man)

C/Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 22-Jun-1999

C/Accession: A53968

R/Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A/Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzym

A/Reference number: A53968; MUID:94308225; PMID:8034709

A/Accession: A53968

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-253 <HAN>

A/Cross-references: GDB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C/Genetics:

A/Genes: GDB:PRSS6; SCCE

A/Cross-references: GDB:377730

A/Map position: 7q35-7q35

C/Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 41.5%; Score 567.5; DB 2; Length 253;  
Best Local Similarity 46.2%; Pred. No. 5.6e-42;  
Matches 111; Conservative 40; Mismatches 82; Indels 7; Gaps 4;

QY 15 LILGVAGSLVSGSCQIINGEDCSPHSQPQWQALWENELFCSGVLVHPQWVLSAHCFCQ 74

Db 16 LALETAGEEAQ--DKIIDGAPCARGHPWQVALLSGNLHCGGVLVNERWVLTAAHCKM 73

QY 75 NSYITGLHSLLEADQEPGQSMVEASLSVRHPEYNRPILLANDMLIKLDESVSSEJTIKS 134

Db 74 NEYTHLGSDTL-GDPR--AQRIKASKFRHPGYSTQTHVNDMLVKNLQARLSMWKK 130

QY 135 ISIASQCTPAGNSCLVSGWGLANG--RMPTVLQCNVSVVSEEVCSKLYDPLIYHPSMPC 192

Db 131 VRLPSRCEPPGTCTVSGWGTTTSPDVTFFPSDLNMCVDVKLISPDQCTKYVQDLLENMLC 190

QY 193 AGGGQDKDSQNGSDGGPLICNGYLGVSFCKAPCGQGVGVVNTLCKPTEWTEKTVO 252

Db 191 AGIPDSKNACNGSDGGPLVCRGTIQLGLVSWGTFFCGQPNDFGVITQVCKFKIWNIDTWK 250

##### RESULT 2

I56559

neuropsin - mouse

C/Species: Mus musculus (house mouse)

C/Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jul-2000

C/Accession: I56559

R/Chen, Z.L.; Yoshida, S.; Kato, K.; Momota, Y.; Suzuki, J.; Tanaka, T.; Ito, J.; Nishin

J. Neurosci. 15, 5088-5097, 1995

A>Title: Expression and activity-dependent changes of a novel limbic-serine protease gen  
A:Reference number: 156559; MUID:95348817; PMID:7623137  
A:Accession: I56559  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-260 <RES>  
A:Cross-references: GB:D30785; NID:g1648847; PIDN:BAA06451.1; PID:g1020091  
C:Superfamily: trypsin; trypsin homology  
F:33-252/Domain: trypsin homology <TRY>

Query Match 41.2%; Score 563.5; DB 2; Length 260;  
Best Local Similarity 41.3%; Pred. No. 1.3e-41;  
Matches 102; Conservative 47; Mismatches 91; Indels 7; Gaps 4;

QY 10 WFLGVLILVAGSLVSGSCSIINGEDCSPHSQPQWQALVWENELFCGVLVHPQWLSA 69  
DB 12 WLLLFPMWAGLFRAGSKLEBECIPESQWQALFOGERLICGVLVGRWVITA 71  
QY 70 AHCPQNSYITGLGHLSEADQEPGQWVEASLSVRHPEYRPL----LANDMLIKLDES 126  
DB 72 AHCKKQKYSVRLGDHSLQSRQP-EQEIQAQSIQHPCYNNSPEDHSHDMLIRLQNSA 130  
QY 127 SESPTIRISASQCPPTAGNSCLVSGGCLLANGR--NPTVLQCVNVSVSEVCSKLYDP 184  
DB 131 NLGDKKVPQVLANLCKFKYQKCIISGWTGTVSPQENFNTLCAEVKIYQNKCCERAYPG 190  
QY 185 LYHPSMFCAGGQDQKDCSCNGSDSGPLICNGYLQGLVSGFKAPCGQVGPVGYTNLCKFT 244  
DB 191 KITEGMVCA-GSSNGADTCQDSSGGFLVCDQGLQISWGSDFGCKEKGVTIKICRYT 249  
QY 245 EWIEKTV 251  
DB 250 TWIKTMT 256

RESULT 3  
A:Title: kallikrein (EC 3.4.21.35) precursor - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 20-Oct-1994 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999  
C:Accession: S45303; S38487  
R:Gauthier, E.R.; Dumas, C.; Chappelain, P.; Tremblay, R.R.; Dube, J.Y.  
Biochim. Biophys. Acta 1218, 102-104, 1994  
A:Title: Characterization of canine pancreas kallikrein cDNA.  
A:Reference number: S45303; MUID:94250683; PMID:8193155  
A:Accession: S45303  
A:Molecule type: mRNA  
A:Residues: 1-261 <GAU>  
A:Cross-references: EMBL:X75479; NID:9414018; PIDN:CAA53210.1; PID:g414019  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-261/Product: tissue kallikrein #status predicted <MAT>  
F:25-253/Domain: trypsin homology <TRY>  
F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 38.3%; Score 524; DB 2; Length 261;  
Best Local Similarity 40.3%; Pred. No. 3.6e-38;  
Matches 106; Conservative 51; Mismatches 84; Indels 22; Gaps 6;

QY 10 WFLGY-LILVAGSLVSGSC---SQIINGEDCSPHSQPQWQALVWENELFCGVLVHPQ 64  
DB 2 WFLVLCALSLAG---TGAAPVQSRIGGWDCTKNSQPWQALYHYSKFCQCGVLVHPE 58  
QY 65 WLSAAHCFQNSYITGLGHLSEADQEPGQWVEASLSVRHPEYRPLAN----- 115  
DB 59 WYVTAACHINDYNQWLGRYNL-FEHEDTAQVQVRESFPHPEFNLSLLKNTLRLPREDY 117  
QY 116 --DLMLIKLDESVDITRISASQCPPTAGNSCLVSGGCLLANGRM--PTVLQCVNV 171  
DB 118 SHDINLLRLAEPAQITDAVRLDLPTQEPQVGSICYSAGWSGIPEFDKFIYDDLLQCV 177  
QY 172 VVSEVCSKLYDPLVHPSMFCAGGQDQKDCSCNGSDSGPLICNGYLQGLVSGFKAPCGQV 231

DB 178 LISNDICANAHSQKVTFFMLCAGHLEGGKDTCTGSGGGLICDGLVQLGITSWGHVPCGSP 237  
QY 232 GVPGVYTNLCKTFETWIEKTVQAS 254  
DB 238 NMPAVYTKVISHLEWIKETMTAN 260

RESULT 4  
A:Title: kallikrein (EC 3.4.21.35) homolog - rat (fragment)  
N:Alternate names: proteinase B light chain  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Nov-1993 #sequence\_revision 18-Nov-1994 #text\_change 22-Jun-1999  
C:Accession: A44284; C41429  
R:Ma, J.X.; Chao, J.; Chao, L.  
Biochemistry 31, 10922-10928, 1992  
A:Title: Molecular cloning and characterization of rKlk10, a cDNA encoding T-kinogenase  
A:Reference number: A44284; MUID:93041794; PMID:1420203  
A:Accession: A44284  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-244 <MAL>  
A:Cross-references: GB:S48142; NID:g259430; PIDN:AAB24071.1; PID:g259431  
A:Experimental source: submandibular gland  
A:Note: sequence extracted from NCBI backbone (NCBIN:118095, NCBI:P:118096)  
R:Kato, H.; Nakanishi, E.; Enjiyoji, K.; Hayashi, S.; Iwanaga, S.  
J. Biochem. 102, 1389-1404, 1987  
A:Title: Characterization of serine proteinases isolated from rat submaxillary gland: wi  
A:Reference number: A41429; MUID:88198057; PMID:3482210  
A:Accession: C41429  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 10-61 <KAT>  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:10-236/Domain: trypsin homology <TRY>  
F:48,103,196/Active site: His, Asp, Ser #status predicted

Query Match 37.6%; Score 515; DB 2; Length 244;  
Best Local Similarity 42.4%; Pred. No. 2e-37;  
Matches 101; Conservative 42; Mismatches 77; Indels 19; Gaps 5;

QY 29 SQIINGEDCSPHSQPQWQALVWENELFCGVLVHPQWLSAAHCFQNSYITGLGHLSEA 88  
DB 8 SRIVGKYCKKNSQPWQVAIL--NEYLCGVLDPSWVITAAHCYSNYHYVLLGRNNLFE 65  
QY 89 DOBPQSQWVEASLSVRHPEYRPL-----ANDMLIKLDESVDITRIS 136  
DB 66 D-EPAQYRFVQNSPPHEDY-KPFLMRNHTQRGDDYNDMLHLSEPADITDGKVID 123  
QY 137 IASQPTAGNSCLVSGWGLL--ANGRMPTVLQCVNVSVSEVCSKLYDPLVHPSMFCAG 194  
DB 124 LPTEEPKVGSTCLASGWSGSKPLANWELPDDLQCVNHLLSNEKCTEAVEQKVTDLMLCAG 183  
QY 195 GQDQDKDCSCNGSDSGPLICNGYLQGLVSGFKAPCGQVGPVYTNLCKTFETWIEKTVQ 252  
DB 184 EMDGRKDTCKGSGGGLICDGLVQLGITSWGNVPCAEYPNPGYTKLIKFTSWIKVMK 241

RESULT 5  
S55066  
A:Title: trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken  
N:Alternate names: trypsinogen II  
C:Species: Gallus gallus (chicken)  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: S55066; S72347  
R:Wang, K.; Gan, L.; Lee, I.; Hood, L.  
Biochem. J. 307, 471-479, 1995  
A:Title: Isolation and characterization of the chicken trypsinogen gene family.  
A:Reference number: S55066; MUID:95251611; PMID:7733885  
A:Accession: S55066  
A:Molecule type: mRNA

A;Title: The amino acid sequence of the gamma-subunit of mouse submaxillary gland 7 S N  
A;Reference number: A92341; MUID:81264363; PMID:7263706  
A;Accession: A92341  
A:Molecule type: protein  
A:Residues: 25-107,112-261 <THO>  
C:Experimental source: outbred strain Swiss Webster  
C:Comment: 7S nerve growth factor is composed of two alpha chains, a beta dimer composed  
C:Comment: The active form of the gamma chain occurs naturally as combinations of either  
C:Genetics:  
A;Map position: 7  
A;Introns: 16/1; 69/2; 165/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: glycoprotein; growth factor; hydrolase; serine proteinase; submandibular gl  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:25-263/Domain: trypsin homology <TRY>  
F:45-107/Domain: segment B1 <GB1>  
F:25-107,112-261/Product: nerve growth factor gamma chain (active form) #status experime  
F:112-261/Domain: segment A <GAA>  
F:112-164/Domain: segment C <GCC>  
F:165-261/Domain: segment B2 <GB2>  
F:31-173,50-66,152-219,184-198,209-234/Disulfide bonds: #status predicted  
F:65,120,213/Active site: His, Asp, Ser #status predicted  
F:102/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 37.2%; Score 509; DB 1; Length 261;  
Best Local Similarity 39.5%; Pred. No. 7.2e-37;  
Matches 103; Conservative 48; Mismatches 86; Indels 24; Gaps 6;

QY 10 WFL-GVLIIGVAG-SLVSGSCQINGEDCSPHSQPQQAALYMENELFCGVLVHPQWVL 67  
Db 2 WFLIIFLAUSLGGIDAAPVQGRIVGGFKCKNNSQPHWVAVRYTGYLGGVLLDPNWVL 61  
QY 68 SAAHQFQNSYITIGLHSLSEADQEPGSQMVSEASLSVRHPPEYRPLL-----AND 116  
Db 62 TAAHQYDDNVKYLWGLKNNLFKD-EPQAQHFVSKAIPHPGFNNLSMRKHIRFLEYDYSND 120  
QY 117 LMLIKDESVSSEDTIRTSISASQCPAGNSCLVSGWGLLANGRMPTVLQ-----CVNV 170  
Db 121 LMLRLSKPADITDTVKPITLTPBPFLGSTCLASGWSIT---PTKQFTDDLYCVNL 176  
QY 171 SVYSSEVCSKLYDPLVHPSMFCAGGQDQKDCSNGDSGGFLICNGYLOGLVSFGKAPCGQ 230  
Db 177 KLLPNEDCAKAIHKYTDAMLCAWEGDGKDTCKGDSGGFLICDGYLOGITSWGHTPCGE 236  
QY 231 VGVPGVYTNLCRPTIEWIKTV 251  
Db 237 PDMPGVYTNLKNFTSWIKDTM 257

RESULT 7  
A34079  
N;Alternate names: kallikrein (EC 3.4.21.35) P1 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Mar-1990 #sequence\_revision 30-Mar-1990 #text\_change 22-Jun-1999  
C:Accession: A34079; S10700  
R;Brady, J.M.; Wines, D.R.; MacDonald, R.J.  
Biochemistry 28, 5203-5210, 1989  
A;Title: Expression of two kallikrein gene family members in the rat prostate.  
A;Reference number: A34079; MUID:89352606; PMID:2765531  
A;Accession: A34079  
A;Status: preliminary  
A:Molecule type: DNA; mRNA  
A:Residues: 1-261 <BRA>  
A;Cross-references: GB:M27215; GB:M27216; GB:M27217; NID:G206638; PIDN:AAA42036.1; PID:G  
A;Experimental source: prostate  
R;Elmoujahed, A.; Gutman, N.; Brillard, M.; Gauthier, F.  
FEBS Lett. 265, 137-140, 1990  
A;Title: Substrate specificity of two kallikrein family gene products isolated from the  
A;Reference number: S10698; MUID:90306305; PMID:2194829  
A;Accession: S10700  
A:Molecule type: protein  
A;Residues: 25-43;112-138 <ELM>

A:Experimental source: submaxillary gland

A>Note: 125-Lys was also found

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:25-253/Domain: trypsin homology <TRY>

F:25-111/Product: tissue kallikrein P1 light chain #status experimental <MAT1>

F:112-261/Product: tissue kallikrein P1 heavy chain #status experimental <MAT2>

F:165,120,213/Active site: His, Asp, Ser #status predicted

Query Match 37.0%; Score 507; DB 2; Length 261;

Best Local Similarity 42.2%; Pred. No. 1.1e-36;

Matches 100; Conservative 39; Mismatches 84; Indels 14; Gaps 3;

QY 29 SOLINGEDCSPHSQPQQAALVMEELFCGVLVHPQWLVAACHFQNSYITGLGLHSLA 88

DB 23 SRIIGGFNCKNSQPVAVVHFNEPQCGGVLIHPSWVITAAHCYSYNYQVWLGRNLL 82

QY 89 DQEPSQMVSEASLSVRHPEYRNPLL-----ANDMLIKLDESVESEDTIRIS 137

DB 83 D-EPPAQHRLVSQSPHFPGFNLDIIKNHTRKPGNDYNDMLLHLKTPADITDGVKVIDL 141

QY 138 ASOCPAGNSCLVSGWGLLA--NGRMPTVLCQVNVSVSEVCSKLYDPLVHPSMFCAGG 195

DB 142 PREEPKVSTCTSGWGSITPLKEFPDDLCQVNIHLLSNEKCIKAYNDEVTDVWLAGE 201

QY 196 GQDKDCNCGDSGGLINGVLQGLVSPGKAPCGQGVPGVYTNLCKTEWIEKTQV 252

DB 202 MDGKDKICKDGGGLICDGLQITSGWSPGEPNKPFSVYTKLIKFTSWMKVMK 258

#### RESULT 8

A27547

trypsin (EC 3.4.21.4) precursor, cationic - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 22-Jun-1999

C:Accession: A27547

R:Fletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.

Biochemistry 26, 3081-3086, 1987

A:Title: Isolation and characterization of a cDNA encoding rat cationic trypsinogen.

A:Reference number: A27547; MUID:87271609; PMID:3607011

A:Accession: A27547

A:Molecule type: mRNA

A:Residues: 1-247 <FLE>

A:Cross-references: GB:M16624; NID:Q206498; PIDN:AAA41985.1; PID:G206499

C:Superfamily: trypsin; trypsin homology

C:Keywords: calcium binding; hydrolase; protein digestion; serine proteinase

F:25-240/Domain: trypsin homology <TRY>

F:31-161,49-65,133-234,140-207,172-186/Disulfide bonds: #status predicted

F:64,108,201/Active site: His, Asp, Ser #status predicted

F:76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 37.0%; Score 506; DB 2; Length 247;

Best Local Similarity 40.5%; Pred. No. 1.2e-36;

Matches 100; Conservative 50; Mismatches 89; Indels 8; Gaps 6;

QY 11 FLGYLLGVAGSL-VSGSCSQINGEDCSPHSQPQQAALVMEELFCGVLVHPQWLVA 69

DB 6 FLAF--LGRAVALPLDDDDKIVGGYTCQKNSLPQVSL-NAGYHFCGSLNSQWVSA 62

QY 70 AHCFQNSYITGLGLHSLAEDQEPGQMVSEASLSVRHPEYRNPILLANDMLIKLDESVESE 129

DB 63 AHCKYGRIQVRLGEHNIDV-VEGGEQFIDAAKIIRHPSYNANTFNDIMLIKNSPATLN 121

QY 130 DTIRSIASQCPAGNSCLVSGWGL--LLANGRMPTVLCQVNVSVSEVCSKLYDPLVH 187

DB 122 SRVSTVSLPRSCGSSGQKLVSGWGLTSLSGTNPSPILLQLCLADPVLSDSCSKSSVPGKIT 181

QY 188 PSMFCAGGQDQDKDCNCGDSGGLICNGYLQGLVSPGKAPCGQGVPGVYTNLCKTEW 247

DB 182 SNMFCLFLGEGKDCSQGDSGSPWCNGLQGVVSWGYG-CAQKGRPGVYTKVNCVNVMI 240

QY 248 EKVQAS 254

:::|:

Db 241 QQTVAAN 247

#### RESULT 9

S01971

tissue kallikrein (EC 3.4.21.35) precursor - mouse

N:Alternate names: glandular kallikrein

C:Species: Mus musculus (house mouse)

C:Date: 01-Dec-1989 #sequence\_revision 01-Dec-1989 #text\_change 21-Jul-2000

C:Accession: S01971; I70023

R:Drinkwater, C.C.; Richards, R.I.

Nucleic Acids Res. 16, 10318, 1988

A:Title: Sequence of mGK-11, a mouse glandular kallikrein gene.

A:Reference number: S01971; MUID:89083511; PMID:3205728

A:Accession: S01971

A:Molecule type: DNA

A:Residues: 1-261 <DRI>

A:Cross-references: EMBL:X13215; NID:G53064; PIDN:CAA31604.3; PID:G5708211

R:Evans, B.A.; Drinkwater, C.C.; Richards, R.I.

J. Biol. Chem. 262, 8027-8034, 1987

A:Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of th

A:Reference number: I55260; MUID:87250386; PMID:3036794

A:Accession: I70023

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 70-122 <RES>

A:Cross-references: GB:M18610; NID:G198529; PIDN:AAA39353.1; PID:G198537

C:Genetics:

A:Introns: 16/1; 69/2; 165/1; 210/3

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-24/Domain: propeptide #status predicted <PRO>

F:25-261/Product: tissue kallikrein #status predicted <MAT>

F:25-253/Domain: trypsin homology <TRY>

F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 37.0%; Score 506; DB 2; Length 261;

Best Local Similarity 40.5%; Pred. No. 1.3e-36;

Matches 104; Conservative 43; Mismatches 94; Indels 16; Gaps 5;

QY 10 WFL-GYLLGVAG-SLVSGSCSQINGEDCSPHSQPQQAALVMEELFCGVLVHPQWL 67

DB 2 WFLIFLALSGIDAAAPPVQSVGFNCFNCKNSQPHWAVRYNKYICGVLLDRNWVL 61

QY 68 SAACHFQNSYITGLGLHSLAEDQEPGQMVSEASLSVRHPEYRNPFL-----AND 116

DB 62 TAACHVSVQNVNWLKTKL-FQEPQAQHRMVSKSPHPDYNMSLLIHNPEDEDSND 120

QY 117 LMLIKLDESVESEDTIRSIASQCPAGNSCLVSGWGLANGRM--PTVLQCVNVSVS 174

DB 121 LMLRLSEPADITDAVKPIALPTEEPKLGSTCLVSGWGSITPTKFTQDDLCQVSKLLP 180

QY 175 EEVCSKLYDPLVHPSMFCAGGQDQKDCNCGDSGGLICNGYLQGLVSPGKAPCGQGVGP 234

DB 181 NEVCVKHNNKQVTDVMLCAGEMGGGKDTCKGDSGGPLICDGVLHGITAWGPIPGKENTP 240

QY 235 GVTNLCCKFTWIEKTIV 251

DB 241 GVTKLKIKFINWIKDTM 257

#### RESULT 10

JE0236

tissue kallikrein (EC 3.4.21.35) - mouse

N:Alternate names: prorenin-converting enzyme

C:Species: Mus musculus (house mouse)

C:Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jul-2000

C:Accession: JE0236

R:Hosoi, K.; Tada, J.; Teamura, K.; Kanamori, N.; Yamanaka, N.

J. Biochem. 124, 369-376, 1998

A:Title: Expression of an allzyme of prorenin-converting enzyme in the submandibular gl

A:Reference number: JE0236; MUID:98351995; PMID:9685728

A;Reference number: S09315; MUID:90225801; PMID:2183721

A;Accession: S09315

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 43-45,'S',47-114,'A',116-261 <BRA>

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; serine proteinase

F;18/Domain: signal sequence #status predicted <SIG>

F;19-261/Product: tissue prokallikrein 7, submandibular #status predicted <MAT>

F;25-253/Domain: trypsin homology <TRY>

F;65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 36.7%; Score 503; DB 2; Length 261;

Best Local Similarity 40.0%; Pred. No. 2.4e-36;

Matches 104; Conservative 42; Mismatches 94; Indels 20; Gaps 5;

QY 10 WFUGY---LILVAGSLVSGSCQIINGEDCSPHSQPQWQAALVWENELFCSGVLVHPQWV 66

DB 2 WFLILFLDLGLGQIDAAPFQ--SRVTGGYKCKNSQWQVALYFTKYLCCGVLIDPSWV 60

QY 67 LSAAHCFQNSYITGLGLHSLEADQPCGSQWVEASLSVRHPEYNRPLL-----A 114

DB 61 ITAHGSSNNYQVWLRNLLLED-EFFAQRILVSQFFPHEDY-KPFLMRNHRKPGDDHS 118

QY 115 NDLMILKLDVSESDETSIRISIASOCPTAGNSCLVSGWGLLAN--GRMPTVLQCVNVSV 172

DB 119 NDLMILHLSQADITGVKVIDLPTBEPKVGSTCLASGWSGTRKPLIWEFFDDLCQVNIHL 178

QY 173 VSEVCSKLYDPLVYHSMFCAGGQKQKSCNGSGGPLICNGYLQGLVSFGKAPCGQGV 232

DB 179 LSNKKCIKAYKEKVTDLMLCAGLEGKGTCTGDSGGPLLCGDLVQLGITSWGSVPCAKTN 238

QY 233 VPGVYTNLCKFTIEWIKTVQ 252

DB 239 MPAIYTKLIKFTSWIKVMK 258

RESULT 12

QJ1472

trypsin (EC 3.4.21.4) V precursor, b-form - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 22-Jun-1999

C;Accession: QJ1472

R;Kang, J.; Wiegand, U.; Mueller-Hill, B.

Gene 110, 181-187, 1992

A;Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.

A;Reference number: QJ1471; MUID:92165057; PMID:1537555

A;Accession: QJ1472

A;Molecule type: mRNA

A;Residues: 1-246 <KAN>

A;Cross-references: EMBL:X59013; NID:957414; PIDN:CAA41752.1; PID:957415

A;Experimental source: pancreas

C;Superfamily: trypsin; trypsin homology

C;Keywords: hydrolase; protein digestion; serine proteinase; zymogen

F;1-15/Domain: signal sequence #status predicted <SIG>

F;16-24/Domain: activation peptide #status predicted <ACT>

F;25-246/Product: trypsin V, b-form #status predicted <MAT>

F;25-239/Domain: trypsin homology <TRY>

F;31-160,49-65,133-233,140-206,171-185/Disulfide bonds: #status predicted

F;64,108,200/Active site: His, Asp, Ser #status predicted

Query Match 36.7%; Score 502.5; DB 2; Length 246;

Best Local Similarity 41.2%; Pred. No. 2.5e-36;

Matches 100; Conservative 47; Mismatches 91; Indels 5; Gaps 5;

QY 14 YLILG-VAGSLVSGSCQIINGEDCSPHSQPQWQAALVWENELFCSGVLVHPQWVLSAAHC 72

DB 7 FTLGLTVAAPFTEDNDRIYGGYTCQBSHPYQVSLNAGSHI-CGGSLLTDQWVLSAAHC 65

QY 73 FQNSYITGLGLHSLEADQPCGSQWVEASLSVRHPEYNRPILLANDIMLIKLDVSESDETI 132

DB 66 YHPQLQVRLGEHNI-YEIEGAEQFIDAAKMLLHLDYDKWTVNDIMLIKLSPATLNSKV 124

[illegible]

RESULT 13  
A29586  
tissue kallikrein (EC 3.4.21.35) hGK-1 precursor - human  
N:Alternate names: glandular kallikrein  
C:Species: Homo sapiens (man)  
C:Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 22-Jun-1999  
C:Accession: A29586  
R:Schedlich, L.J.; Bennetts, B.H.; Morris, B.J.  
DNA 6, 429-437, 1987  
A:Title: Primary structure of a human glandular kallikrein gene.  
A:Reference number: A29586; MUID:8805467; PMID:2824146  
A:Accession: A29586  
A:Molecule type: DNA  
A:Residues: 1-261 <SCH>  
A:Cross-references: GB:M18157; NID:G186640; PIDN:AAA74454.1; PID:G386942  
A:Note: the authors translated the codon TAC for residue 43 as TTP  
C:Generics:

C:genetics: 16/1; 69/2; 165/1; 210/3  
A;introns: 16/1; 69/2; 165/1; 210/3  
C:superfamily: trypsin; trypsin homology  
C:keywords: hydrolase; serine proteinase  
F;25-253/domain: trypsin homology <TRY>  
F;65;120;213/Active site: His, Asp, Ser #status predicted  
Query March 36.7%; Score 502.5; DB 2; Length 261;  
Best Local Similarity 38.7%; Pred. No. 2.7e-36;  
Matches 101; Conservative 44; Mismatches 91;  
Indels 25; Gaps 5;

	Qy	15	LILGVAGSLVSGC-----SQINGEDCSPHSOPWQAALVMENELFCSGVLVHPQWY 66    :   :::
Dd	Db	4	LVLZIA---LSVGCTGAUPLIQSRIVGWGCEBKHSQPFWAVYSHGWAHCGGYLVHPQWY 60    :   :::
	Qy	67	LSAAHFQNFSVTTCGLGHSTLEADQPCSSOMVBASLVRHPEYNRPLL-----AN 115    :   :::
Dd	Db	61	LTRAHLCKKNSQWLGRHNLFEPEDTG-QRVFVSJSHFPHELNYMSLLKHQS LAPDEDSH 119    :   :::
	Qy	116	DLMLIKIDESVESDTSIRISIAQCPTAGNCSCLVSGWGLANGRM--PVLQCVNVSVV 173    :   :::
Dd	Db	120	DLMLRLSEPAKTIDVVVKVLGTQEAPALTGCYASCWGSIETEEFLRPSLRSCQVSLIHL 179    :   :::
	Qy	174	SBEVCSKLYDFLPHSPMCFACGGGDOKDCNGSDSGGPLICNGYLQGIVSFPGKAPCCQGV 233    :   :::
Dd	Db	180	SNDMCARAYSEKVTEFMLCAGLTGGDKTCGGSDSGGPLVCNCVLOQTITSWGPFECALPEX 239    :   :::
	Qy	234	PgyvtnlckfIEWIEkTVqAs 254    :   :::
Dd	Db	240	pAvytkvvhvkwtKDtiAAn 260    :   :::

RESULT 14

A41020 tissue kallikrein (EC 3.4.21.35) mGK-13 precursor - mouse

N:Alternate names: epidermal growth factor-binding protein type B; prorenin-converting enzyme

C:Species: Mus musculus (house mouse)

C:Date: 13-May-1992 #sequence revision 13-May-1999 #text\_change 22-Jun-1999

C:Accession: A41020; B29746; F02014; I70027; S18674

E:Klm, W.S.; Nakayama, K.; Nakagawa, T.; Kawamura, Y.; Haraguchi, K.; Murakami, K.

J. Biol. Chem. 266, 19283-19287, 1991

A:Title: Mouse submandibular gland prorenin-converting enzyme is a member of glandular k

A:Reference number: A41020; MUID:92011720; PMID:1918045

A:Accession: A41020

A;Molecule type: mRNA  
A;Residues: 1-261 <IM>  
A;Cross-references: GB:X58628; NID:953771; PIDN:CAAA1482.1; PID:953772  
R;Drinkwater, C.C.; Evans, B.A.; Richards, R.I.  
Biochemistry 26, 6750-6756, 1987  
A;Title: Mouse glandular kallikrein genes: identification and characterization of the  $\alpha$  1 gene  
A;Reference number: A90522; M0ID:88107594; PMID:3322387  
A;Accession: B29746  
A;Molecule type: DNA  
A;Residues: 1-118, 'D', 120-261 <DR1>  
A;Experimental source: Strain BALB/c, salivary gland  
K;Hosoi, K.; Tsunawasa, S.; Kurihar, K.; Aoyama, H.; Ueha, T.; Murai, T.; Sakiyama, F.  
J. Biochem. 115, 137-143, 1994  
A;Title: Identification of mk1, a true tissue (glandular) kallikrein of mouse submandibular gland

B. A:Reference number: PC2013; MUID:94245648; PMID:8188620  
A:Accession: PC2014  
A:Molecule type: protein  
A:Residues: 25-58, 'X', 60-66; 165-206 <HOS>  
A:Experimental source: submandibular gland  
A:Evans, B.A.; Drinkwater, C.C.; Richards, R.I.  
J. Biol. Chem. 262, 8037-8034, 1987  
A:Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of t  
A:Reference number: I55260; MUID:87250386; PMID:3036794  
A:Accession: I70027  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 70-132 <RES>  
A:Cross-references: GB:M18612; NID:gl198533; PIDN:AAA39354.1; PID:gl198538  
C:Genetics:  
A:Gene: mk13  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; saliva; serine proteinase; submandibular gland  
F:25-253/Domain: trypsin homology <TRY>  
F:65-120, 213/Active site: His, Asp, Ser #status predicted

Query Match	36.7%	Score	502;	DB 2;	Length	261;			
Best Local Similarity	40.1%;	Pred. No.	2.9e-36;						
Matches	103;	Conservative	44;	Mismatches	94;	Indels	16;	Gaps	5;
Qy	10	WFL-GVILGVAG-SLVGSGCSOIINGEDCSPHSQPQAAALVMENELFCSGVLVHPQWVL	67						
Db	2	WFLILFLALSUGIDAAPPLQSRVVGFCNKSQPMQVAVYQKEHICGVLLDRNWVL	61						
Qy	68	SAAHCFQNSYTIIGLHLSLEADQPGSQMVEASLSVRHPEYNRPLL-----AND	116						
Db	62	TAAHCYVDQYEVLGKNKL-FQEESPQAHLVSKSPHPGFNMSLLMLQTIIPGADFSND	120						
Qy	117	LMILKIDESVSEPTIRISIASQCPTAGNSCLVSGWGLLANGR--MPTVLQCVNVSVVS	174						
Db	121	LMLLRLSKPADITDVVKPIALPTKEPKPGSKCLASGWSITPTRWQFPDDLQCVFETLLP	180						
Qy	175	EEVCSKLYDLYHPSMFCAGGGQDKDSNCGDSGGPLICNGYLQGLVSYFGKAPCGQGVCP	234						
Db	181	NENCAVYLOKVTDVMLCAGMGGGKDTCRDDSGGPLICDIBLOQTTSYGPVPCGKGVCP	240						
Qy	235	GVYTNLCKFTIEWIEKTV	251						
Db	241	AIYTNLIKFNISWKDTM	257						

RESULT 15  
T01779  
trypsin (EC 3.4.21.4) - plaice  
C:Species: pleuronectes platessa (plaice)  
C;Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 16-Jul-1999  
C;Accession: T01779  
R;Leaver, M.J.; George, S.G.  
submitted to the EMBL Data Library, March 1996  
A;Reference number: Z14422  
A;Accession: T01779  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA

A:Residues: 1-250 <LEA>  
A:Cross-references: EMBL:X56744; NID:G1213630; PID:G64240  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:23-242/Domain: trypsin homology <TRY>

Query Match 36.6%; Score 501.5; DB 2; Length 250;  
Best Local Similarity 42.3%; Pred. No. 3.1e-36;  
Matches 105; Conservative 37; Mismatches 97; Indels 9; Gaps 5;

QY	12	LGVLILGVAGSLVSGSCQIINGEDCSPHSQPQAAALVMENELFCSGVLVHPQWVLSAAH	71
DB	4	LALLLVGAAGAVVPREDGRIIGGHECAHSRPFMASLNYGYH-FCGGVLIINNQWVLSVAH	62
QY	72	CFONSYT--IGLGLHSLKADQEPGSCQWVEASLSVRHPEYNRPILLANDLMLIKLDESVSSES	129
DB	63	CWNPYAMQVNLGEHLR-KFEGTEQLMKTDITIIHPNIDYQTLDFDINLILKLYHPVEVS	121
QY	130	DTIRISIASQCPTAGNSCLVSGWGLLANG---RMPTVLQCVNVSWSVEEVCCKLYDPL	185
DB	122	HAVGPIPLPTSCPVAGTPCSVSGWGNRTARDGDEVLPITLLQCMDFIVDDEECMKSYDPM	181
QY	186	YHPSMECAGGGQDKSCNGDSGPLICNGYLGKAFQCGQGVGVGVYTNLCKPTE	245
DB	182	ISPRMYCAGFMDSGRDAGCNGDSGPLVCRGEVYGLVSWGQG-CAQPNYPGVYVYKLCEFLG	240
QY	246	WIEKTVQA	253
DB	241	WIERLEA	248

Search completed: May 5, 2004, 16:09:56  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 16:04:07 ; Search time 45 Seconds  
(without alignments)  
1780.925 Million cell updates/sec

Title: US-09-895-814-525  
Perfect score: 1369  
Sequence: 1 NATAGPMWFLGLYLGLVA.....GVYTNLCKFTWIKTVQAS 254

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:  
1: sp archaea:  
2: sp bacteria:  
3: sp fungi:  
4: sp human:  
5: sp invertebrate:  
6: sp mammal:  
7: sp mhc:  
8: sp organelle:  
9: sp phase:  
10: sp plant:  
11: sp rodent:  
12: sp virus:  
13: sp vertebrate:  
14: sp unclassified:  
15: sp rvirus:  
16: sp bacteriaph:  
17: sp archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1113	81.3	254	6 Q9XSN6	Q9XSN6 sus scrofa
2	1098	80.2	205	4 Q96JEO	Q96JEO homo sapien
3	1087	79.4	205	4 Q96JED	Q96JED homo sapien
4	1089.5	78.1	204	4 Q96JED	Q96JED homo sapien
5	1010	73.8	255	11 Q9JIS2	Q9JIS2 mus musculus
6	995	72.7	255	11 Q9ZOM1	Q9ZOM1 mus musculus
7	848	61.9	195	4 Q96PT0	Q96PT0 homo sapien
8	845	61.7	159	4 Q96PT0	Q96PT0 homo sapien
9	675	49.3	131	4 Q96JED	Q96JED homo sapien
10	637.5	46.6	293	11 Q9D140	Q9D140 mus musculus
11	577	42.1	146	4 Q96JED	Q96JED homo sapien
12	575	42.0	269	4 Q9IUS5	Q9IUS5 homo sapien
13	574	41.9	110	4 Q96RUS	Q96RUS homo sapien
14	556.5	40.7	253	4 Q8NSN9	Q8NSN9 homo sapien
15	540.5	39.5	242	13 Q7T1R8	Q7T1R8 pangasius h
16	538.5	39.3	250	11 Q8CGR5	Q8CGR5 mus musculus

17	537.5	39.3	242	11	Q80VS4	Q80VS4 mus musculus
18	537	39.2	276	11	Q8CGR6	Q8CGR6 mus musculus
19	529.5	38.7	275	4	Q8IXD7	Q8IXD7 homo sapien
20	529	38.6	249	11	Q91VE3	Q91VE3 mus musculus
21	524	38.3	249	11	Q9QXN4	Q9QXN4 mus musculus
22	524	38.3	261	6	Q29474	Q29474 canis famil
23	524	38.3	276	11	Q9QXN3	Q9QXN3 m hipposcas
24	517	37.8	247	11	Q9CPN9	Q9CPN9 mus musculus
25	516	37.7	235	11	Q63274	Q63274 rattus norv
26	515	37.6	249	13	Q92046	Q92046 dissostichu
27	513.5	37.5	260	4	Q8IW69	Q8IW69 homo sapien
28	512	37.4	239	11	Q63275	Q63275 rattus norv
29	509	37.2	261	6	Q9N1Q1	Q9N1Q1 saguinus oe
30	507.5	37.1	247	11	Q9D7Y7	Q9D7Y7 mus musculus
31	507	37.0	254	11	Q8CGR4	Q8CGR4 mus musculus
32	506	37.0	247	11	Q9CPN7	Q9CPN7 mus musculus
33	505.5	36.9	250	13	Q93265	Q93265 pseudopleur
34	504	36.8	249	13	Q9W6K0	Q9W6K0 notothenia
35	503	36.7	261	11	Q88309	Q88309 mus musculus
36	503	36.7	675	13	Q9W6J8	Q9W6J8 dissostichu
37	498.5	36.4	246	11	Q7T142	Q7T142 mus musculus
38	498	36.4	286	4	Q8ETG8	Q8ETG8 homo sapien
39	496.5	36.3	255	4	Q96RQ0	Q96RQ0 homo sapien
40	493	36.0	248	13	Q7S2T1	Q7S2T1 xenopus lae
41	492.5	36.0	251	11	O54854	O54854 rattus norv
42	492	35.9	263	11	Q9JW69	Q9JW69 mus musculus
43	489.5	35.8	262	4	Q86U61	Q86U61 homo sapien
44	486.5	35.5	261	11	Q9JW70	Q9JW70 mus musculus
45	485	35.4	234	11	Q9R048	Q9R048 mus musculus

ALIGNMENTS

RESULT 1

Q9XSN6 PRELIMINARY; PRT; 254 AA.  
AC Q9XSN6;  
ID Q9XSN6;  
DT 01-NOV-1999 (TRENBLrel. 12, Created)  
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Enamel matrix serine proteinase 1 precursor.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98126310; PubMed=9465170;  
RA Simmer J.P., Fukae M., Tanabe T., Yamakoshi Y., Uchida T., Xue J.,  
RA Margolis H.C., Shimizu M., DeHart B.C., Hu C.-C., Bartlett J.D.;  
RT "Purification, characterization, and cloning of enamel matrix serine  
RT proteinase 1.";  
RL J. Dent. Res. 77:377-386(1998).  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; U76256; AAB94638.1; -.  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.251; -.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004235; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR01314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_Ser; 1.  
DR PROSITE; PS00240; TRYP\_SIN\_DOM; 1.  
DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
DR Q8NSN9; Q8NSN9; 1.  
KW Hydroxylase; Protease; Serine protease; Signal.  
FT SIGNAL 1 24







01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
Enamel matrix serine proteinase 1 precursor.  
KLK4.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
J. Dent. Res. 79:70-76(2000).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=Swiss-Webster;  
MEDLINE=20152522; PubMed=10690663;  
Hu J.C., Ryu O.H., Chen J.J., Uchida T., Wakida K., Murakami C.,  
Jiang H., Qian Q., Zhang C., Otmers V., Bartlett J.D., Simmer J.P.;  
"Localization of EMSP1 expression during tooth formation and cloning  
of mouse cDNA.";  
in: developing mouse incisors.";  
Gene 251:1-8(2000).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=Swiss-Webster;  
MEDLINE=20323211; PubMed=10863090;  
Hu J.C., Zhang C., Sun X., Yang Y., Cao X., Ryu O., Simmer J.P.;  
"Characterization of the mouse and human PRSS17 genes, their  
relationship to other serine proteases, and the expression of PRSS17  
in developing mouse incisors.";  
Gene 251:1-8(2000).  
[2]  
-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
EMBL; AF019979; AAC98894.1; -.  
HSSP; P00763; 1DPO.  
MEROPS; S01.251; -.  
MGD; MGI:1861379; KLK4.  
GO; GO:0004263; F:chymotrypsin activity; IEA.  
GO; GO:0008233; F:peptidase activity; IEA.  
GO; GO:0004295; F:trypsin activity; IEA.  
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
InterPro; IPR009003; Cys\_Ser\_trypsin.  
InterPro; IPR001254; Peptidase\_S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00089; trypsin\_1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
Hydrolase; Protease; Serine protease; Signal.  
SIGNAL  
CHAIN 1 25 MATRIX SERINE PROTEINASE 1.  
SEQUENCE 255 AA; 27488 MW; 6FD2E7DEA0660A2A CRC64;  
Query Match 72.7%; Score 995; DB 11; Length 255;  
Best Local Similarity 69.3%; Pred. NO. 1e-87;  
Matches 176; Conservative 31; Mismatches 47; Indels 0; Gaps 0;  
QY 1 MATAGNPGWFLGVLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
DB 2 MVTARTPGWFLGVLILEVTGASASSVSSRIIQQDCSPHSQPWOAALFSGEDFCGVL 61  
QY 61 VHPQWLSAAHCFONSYYTIGLHSLSEADQEPGQSQVSEASLSVRHPYRNPLLANDLMLI 120  
DB 62 VHPQWLSAAHCLQESYIVGLHNLKGSQPGSRMLEAHLISIQHPNFDPSFANDLMLI 121  
QY 121 KLDESVSDDTIRSIASQCPAGNSCLVSGWGLANGMPTVLCQVNVSVSEVCSK 180  
DB 122 KLVESVIESNTRISFVATQCTPGTCLVSGWGLKNGKLPSSLQCVNLVSASEEFCRL 181  
QY 181 LYDPLYPHMFAGGQDQKDCNDSGGPLICNGYQLGVSFGKAPCGQGVGVVYTNL 240  
DB 182 LYDPVYHLMFAGGQDQKDCNDSGGPIVCNRSQGLVSMGQKCGQGPVSVYTNL 241  
QY 241 CKFTWIKTVQAS 254  
DB 242 CKFTWIKTVQAS 255

RESULT 7  
Q96PTO PRELIMINARY; PRT; 195 AA.  
AC Q96PTO;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Kallikrein 4 splice variant.  
DE KLK4.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99367447; PubMed=10438493;  
RA Stephenson S.A.; Verity K.; Ashworth L.K.; Clements J.A.;  
"Localization of a new prostate-specific antigen-related serine  
protease gene, KLK4, is evidence for an expanded human kallikrein gene  
family cluster on chromosome 19q13.3-13.4.";  
J. Biol. Chem. 274:23210-23214(1999).  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; AF148532; AAL14782.1; -.  
DR HSSP; P00761; 1ANL.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00089; trypsin\_1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
Hydrolase; Protease; Serine protease.  
KW SEQUENCE 195 AA; 20585 MW; EF7E96978B1515B1 CRC64;  
Query Match 61.9%; Score 848; DB 4; Length 195;  
Best Local Similarity 97.6%; Pred. NO. 1.1e-73;  
Matches 160; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATAGNPGWFLGVLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
DB 1 MATAGNPGWFLGVLILGVAGSLVSGSCQIINGEDCSPHSQPWOAALVMENELFCSGVL 60  
QY 61 VHPQWLSAAHCFONSYYTIGLHSLSEADQEPGQSQVSEASLSVRHPYRNPLLANDLMLI 120  
DB 62 VHPQWLSAAHCFONSYYTIGLHSLSEADQEPGQSQVSEASLSVRHPYRNPLLANDLMLI 120  
QY 121 KLDESVSDDTIRSIASQCPAGNSCLVSGWGLANGMPTV 164  
DB 121 KLDESVSDDTIRSIASQCPAGNSCLVSGWGLANGELTGV 164  
RESULT 8  
Q96PTI PRELIMINARY; PRT; 159 AA.  
AC Q96PTI;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Kallikrein 4 splice variant.  
DE KLK4.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99367447; PubMed=10438493;  
RA Stephenson S.A.; Verity K.; Ashworth L.K.; Clements J.A.;

"Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3-13.4.";  
J. Biol. Chem. 274:23210-23214 (1999).  
-/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; AF148532; AAL14781.1; -  
DR HSP; P00761; 1AN1.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin.1  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 159 AA; 16931 MW; 3C66BFA91B6CA5A CRC64;  
  
Query Match 61.7%; Score 845; DB 4; Length 159;  
Best Local Similarity 100.0%; Pred. No. 1.7e-73;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MATAGNPNWFLGYLLGVAGSLVSGSCQIINGDCSPHSQPWQAALVNMENLFCSGVL 60  
DB 1 MATAGNPNWFLGYLLGVAGSLVSGSCQIINGDCSPHSQPWQAALVNMENLFCSGVL 60  
QY 61 VHPQWVLSAHCFCQNSVTIGLHSLHSLADQEPGQVMSVSRHPEYRPLANDMLLI 120  
DB 61 VHPQWVLSAHCFCQNSVTIGLHSLHSLADQEPGQVMSVSRHPEYRPLANDMLLI 120  
QY 121 KLDSVESDSTIRISIASOCPAGNSCLVSGWGLLANG 159  
DB 121 KLDSVESDSTIRISIASOCPAGNSCLVSGWGLLANG 159  
  
RESULT 9  
Q96JDB PRELIMINARY; PRT; 131 AA.  
ID Q96JDB  
AC Q96JDB  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE ARM1 (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX MEDLINE=21398046; PubMed=11506707;  
RA Korkmaz K.S., Korkmaz C.G., Pretlow T.G., Saatcioglu F.;  
RT "Distinctly different gene structure of KLK4/CLK-11/Protease/ARM1 compared with other members of the kallikrein family - Perinuclear localization, alternative cDNA forms and regulation by multiple hormones.";  
RT DNA Cell Biol. 20:435-445(2001).  
RL -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC EMBL; AF259968; AAK71704.1; -  
DR HSP; P00761; 1AN1.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin.1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.

SMART; SM00020; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 131 131  
SQ SEQUENCE 131 AA; 14107 MW; 03434B6D95AF2406 CRC64;  
  
Query Match 49.3%; Score 675; DB 4; Length 131;  
Best Local Similarity 100.0%; Pred. No. 3.3e-57;  
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 50 MENELFCGVLVHPQWVLSAHCFCQNSVTIGLHSLHSLADQEPGQVMSVSRHPEYR 109  
DB 1 MENELFCGVLVHPQWVLSAHCFCQNSVTIGLHSLHSLADQEPGQVMSVSRHPEYR 60  
QY 110 RPLANDMLLIKLDSVESDSTIRISIASOCPAGNSCLVSGWGLLANGRMTVLQCVN 169  
DB 61 RPLANDMLLIKLDSVESDSTIRISIASOCPAGNSCLVSGWGLLANGRMTVLQCVN 120  
QY 170 VSVVSEEVCS 179  
DB 121 VSVVSEEVCS 130  
  
RESULT 10  
Q9DI40 PRELIMINARY; PRT; 293 AA.  
ID Q9DI40  
AC Q9DI40  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE 1110030019Rik protein.  
GN 1110030019RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,  
RA Saio T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Marzari J., Monbets P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Yashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RT Nature 409:685-690(2001).  
RL -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC EMBL; AK003996; BAB23113.1; -  
DR HSP; P00763; 1DPO.  
DR MEROPS; S01.017; -  
DR MEROPS; S01.418; -  
DR MGD; MGI:131518; 1110030019Rik.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.



```
Db 181 AGTKCLVSGGTTKSPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTWFCAGDKRGRD 239
Qy 202 SCNGDSGGLICNGYLGQVLSFGKAPCGQ 230
Db 240 SCQDSGPGVVCNGSLQGLVSGWDYPCAR 268

RESULT 13
Q96RUS PRELIMINARY; PRT; 110 AA.
AC Q96RUS;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ARMI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21398046; PubMed=11506707;
RA Korkmaz K.S., Korkmaz C.G., Prellow T.G., Saatcioglu F.;
RT "Distinctly different gene structure of KLK4/CLK-L1/protease/ARMI
RT compared with other members of the kallikrein family - Perinuclear
RT localization, alternative cDNA forms and regulation by multiple
RT hormones.";
RL DNA Cell Biol. 20:435-445(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF259971; AAK71706.1; -.
DR HSP; P00761; IAN1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004395; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 110 AA; 11858 MW; B6F9C135EA93B116 CRC64;

Query Match 41.9%; Score 574; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 MENELFCGVLVHPQWVLSAAHCFQNSYTTGLGLHSLEADQEPGQWVEASLSVRHPEYN 109
Db 1 MENELFCGVLVHPQWVLSAAHCFQNSYTTGLGLHSLEADQEPGQWVEASLSVRHPEYN 60

Qy 110 RPLANDMLIKLDESVSSEDTRTSISIASQCTAGNSCLVSGWGLANG 159
Db 61 RPLANDMLIKLDESVSSEDTRTSISIASQCTAGNSCLVSGWGLANG 110

RESULT 14
Q8EN59 PRELIMINARY; PRT; 253 AA.
AC Q8EN59;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kallikrein 7 (chymotryptic, stratum corneum).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC032005; AAH32005.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 40.7%; Score 556.5; DB 4; Length 253;
Best Local Similarity 45.8%; Pred. No. 2e-45;
Matches 110; Conservative 40; Mismatches 83; Indels 7; Gaps 4;

Qy 15 LILVAGSLVSGSCSQIINGEDCSPHSQPQWQAALVMENELFCSGLVHPQWVLSAAHCFQ 74
Db 16 LALETAGEAQG--DKIIDGAPCARGHPQWQVALLSGNQLHCGGVLVNERWLTAAACRM 73

Qy 75 NSYITGLHLSLEADQEPGQWVEASLSVRHPEYNRPILLANDMLIKLDESVSSEDTRTS 134
Db 74 NEYIVHLGSDTL-GDRL--AQIRAKSRFRHPGYSYQTHVNDMLVKNLSQARLSMVKK 130

Qy 135 ISIASQCTAGNSCLVSGWGLANG--RMPTVLCQNVSVVSEVCSKLYDPLVHPSMFC 192
Db 131 VRLPSRCPEPPTCTCTVSGWGTTFDPVTFPSDLACVDVKLISPDQCTKYKDLLENMLC 190

Qy 193 AGGGQDQKDCNCGSGGGLICNGYLGQVLSFGKAPCGQWVGVVYTNLCKTEMIETVQ 252
Db 191 AGIPDSKKNACNGSGGGLVCRGLQGLVSGWTFPGWQPDGYYTQVCKTKINDMK 250

RESULT 15
Q7TIR8 PRELIMINARY; PRT; 242 AA.
AC Q7TIR8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Trypsinogen.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Pangasiidae; Pangasius.
OX NCBI_TaxID=85560;
RN [1]
RP SEQUENCE FROM N.A.
RA Moreau Y., Avarre J.-C., Cahu C., Suryanti Y., Utami R.;
RT "Characterization of trypsin and cDNA encoding trypsinogen in a
RT pangasiidae, Pangasius hypophthalmus.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY316360; AAP81159.1; -.
DR SEQUENCE 242 AA; 26361 MW; 17D9B1F71B08B5306 CRC64;

Query Match 39.5%; Score 540.5; DB 13; Length 242;
Best Local Similarity 42.4%; Pred. No. 6.6e-44;
Matches 101; Conservative 54; Mismatches 74; Indels 9; Gaps 5;

Qy 23 LVSSC-----SQIINGEDCSPHSQPQWQAALVMENELFCSGLVHPQWVLSAAHCFQNSY 77
Db 8 LLVGACFALEDKDKVGGVECTYSPQWQVSLNVGYH-FCGSLINQNQNWWSAAHCFQNSRI 66
```

7

Qy	78	TIGLGLSLEADQPGSGWYEA	SLSVHPBPWEPLLANDMLIKDSESDTTRTSI	133
:	:	:	:	:
:	:	:	:	:
:	:	:	:	:
Db	67	EVRLEGHIQIN-EQTGFIS	RVPNPNYSWTINDIMLIKLSQSASVNNYQPVAL	125
:	:	:	:	:
:	:	:	:	:
:	:	:	:	:
Qy	138	ASQCETAGNSCLVSGWG-L	IANGRMPTVLQCNVSVSEVCSKLYDPHYHSMRCAGG	196
:	:	:	:	:
:	:	:	:	:
Db	126	PSSCFPAGTWICVSGWN	TNESTADRNLKCVEPIILSDCCNNSTFGMTTKAMCAEFL	185
:	:	:	:	:
:	:	:	:	:
Qy	197	QDKDSKNGDSGGPLINC	GYLGIVSFGEKAPCGGVGVFWYTNLCFTBWIERTVOAS	254
:	:	:	:	:
:	:	:	:	:
Db	186	EGGKDCSCDGSGGPVC	NELQGVISNGYG-CAEKHFVGYYTKVICTFDWAQTASN	242
:	:	:	:	:
:	:	:	:	:

Search completed: May 5, 2004, 16:09:24  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 15:55:17 ; Search time 17 Seconds  
(without alignments)  
777.989 Million cell updates/sec

Title: US-09-895-814-525

Perfect score: 1369

Sequence: 1 MATAGNFWGFLVGLGVA.....GVYNLCKFTWIEKTVQAS 254

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1369	100.0	254	1 KLK4_HUMAN	Q9Y5K2 homo sapien
2	671	49.0	293	1 KLK2_MOUSE	O95237 homo sapien
3	569.5	41.6	260	1 NRPN_RAT	P00756 mus musculus
4	567.5	41.5	253	1 KLK7_HUMAN	P49862 homo sapien
5	563.5	41.2	260	1 NRPN_MOUSE	P49862 mus musculus
6	549	40.1	250	1 KLK8_HUMAN	Q9UBX7 homo sapien
7	535.5	39.1	251	1 KLK6_HUMAN	Q9P0G3 homo sapien
8	517.5	37.8	277	1 KLK3_HUMAN	Q9UKR3 homo sapien
9	515.5	37.7	260	1 KLK8_HUMAN	O60259 homo sapien
10	515	37.6	244	1 KLK4_RAT	P36375 rattus norv
11	510.5	37.3	248	1 TRY3_CHICK	Q90629 gallus gall
12	509	37.2	261	1 KLK3_MOUSE	P00756 mus musculus
13	507	37.0	261	1 KLK8_RAT	P36374 rattus norv
14	506.5	37.0	244	1 KLK6_HUMAN	Q92876 homo sapien
15	506	37.0	247	1 TRY3_RAT	P08426 rattus norv
16	506	37.0	261	1 KLK8_MOUSE	P15946 mus musculus
17	503	36.7	261	1 KLK7_RAT	P36373 rattus norv
18	502.5	36.7	246	1 TRY2_HUMAN	P32822 rattus norv
19	502.5	36.7	261	1 KLK2_HUMAN	P20151 homo sapien
20	502	36.5	261	1 KLK3_MOUSE	P36368 mus musculus
21	499.5	36.5	246	1 TRY1_CANFA	P06871 canis fami
22	498	36.4	248	1 KLK6_HUMAN	Q9UKR0 homo sapien
23	498	36.4	256	1 KLK6_HUMAN	Q9H2R5 homo sapien
24	498	36.4	261	1 KLK3_HUMAN	P07288 homo sapien
25	497	36.3	247	1 TRY4_RAT	P12788 rattus norv
26	497	36.3	259	1 KLK3_RAT	P36376 rattus norv
27	496	36.2	248	1 TRY2_CHICK	Q90628 gallus gall
28	496	36.2	259	1 KLK3_MOUSE	P15948 mus musculus
29	496	36.2	261	1 KLK1_RAT	P00758 rattus norv
30	495.5	36.2	248	1 TRY1_RAT	P32821 rattus norv
31	493.5	36.0	248	1 TRY1_CHICK	Q90627 gallus gall
32	493	36.0	244	1 TRY2_XENLA	P70059 xenopus lae
33	492.5	36.0	261	1 KLK3_MOUSE	P15949 mus musculus

#### ALIGNMENTS

##### RESULT 1

ID	KLK4_HUMAN	STANDARD;	PRT;	254 AA.
AC	Q9Y5K2; Q9GZL6; Q9UBJ6;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Kallikrein 4 precursor (EC 3.4.21.-) (protease) (Kallikrein-like protein 1) (KLK-L1) (Enamel matrix serine proteinase 1).			
GN	KLK4 OR PRS17 OR PSTS OR ENSPI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99179024; PubMed=10077646;			
RA	Nelson P.S., Gan L., Ferguson C., Moss P., Gelinas R., Hood L., Wang K.,			
RA	"Molecular cloning and characterization of prostate, an androgen-regulated serine protease with prostate-restricted expression.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 96:3114-3119(1999).			
LN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99413477; PubMed=10485467;			
RA	Yousef G.M., Obezu C.V., Luo L.-Y., Black M.H., Diamandis E.P.;			
RT	"Protease/KLK-L1 is a new member of the human kallikrein gene family, is expressed in prostate and breast tissues, and is hormonally regulated.";			
RL	Cancer Res. 59:4252-4256(1999).			
LN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99367447; PubMed=10438493;			
RA	Stephenson S.A., Verity K., Ashworth L.K., Clements J.A.;			
RT	"Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3-13.4.";			
RL	J. Biol. Chem. 274:23210-23214(1999).			
LN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20510030; PubMed=11054574;			
RA	Gan L., Lee I., Smith K., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepser B., Wang K.;			
RT	"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";			
RL	Gene 257:119-130(2000).			
LN	[5]			
RP	SEQUENCE FROM N.A.			
RX	Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J., Liu S., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreau T., Frankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B., Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;			

Q07276 macaca fasc  
P36369 mus musculus  
P15947 mus musculus  
P00760 bos taurus  
Q28773 papio hamad  
P06870 homo sapien  
P00762 rattus norv  
P33619 macaca mlla  
P00759 rattus norv  
P15945 mus musculus  
Q29463 bos taurus  
P00755 mus musculus

34 492 35.9 257 1 KLK1\_MACPA  
35 491 35.9 261 1 KLK2\_MOUSE  
36 489.5 35.8 261 1 KLK6\_MOUSE  
37 488.5 35.7 243 1 TRY1\_BOVIN  
38 488.5 35.7 258 1 KLK1\_PAPHA  
39 487.5 35.6 262 1 KLK1\_HUMAN  
40 485 35.4 246 1 TRY1\_RAT  
41 485 35.4 261 1 KLK3\_MACMU  
42 484 35.4 259 1 KLK2\_RAT  
43 484 35.4 261 1 KLK5\_MOUSE  
44 480 35.1 247 1 TRY2\_BOVIN  
45 480 35.1 261 1 KLK1\_MOUSE



"Sequence analysis of a 4.8 MB region of 19q13.4 between KLK1 and SYT3.";  
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE OF 22-254 FROM N.A.  
 RA Simmer J.P., Ryu O.H., Qian Q., Zhang C., Cao X., Sun X., Hu C.-C.;  
 RT "Cloning and characterization of a cDNA encoding human EMSPI.";  
 RL (In) Goldberg M. (eds.);  
 RL Chemistry and biology of mineralized tissues, pp.1-1, American  
 RL Academy of Orthopaedic Surgeons, Vittel (2000).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed in prostate.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. kallikrein subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF113140; RAD21580.1; -;  
 CC EMBL; AF113141; AAD21581.1; -;  
 CC EMBL; AF135023; AAD26424.2; -;  
 CC EMBL; AF148532; AAD38019.1; -;  
 CC EMBL; AF243527; AAG33357.1; -;  
 CC EMBL; AC037199; -; NOT ANNOTATED\_CDS.  
 CC EMBL; AF126401; AAG43246.1; -;  
 CC HSSP; P00763; LDPO.  
 CC MEROPS; S01.251; -;  
 CC Genew; HGNC:6365; KLK4.  
 CC MIM; 603767; -;  
 CC GO; GO:0005615; C:extracellular space; TAS.  
 CC GO; GO:000236; F:serine-type peptidase activity; TAS.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR001254; Peptidase S1.  
 CC InterPro; IPR001314; Peptidase\_S1a.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Hydrolase; Serine protease; Zymogen; Signal.  
 KW SIGNAL 1 26  
 FT PROPEP 27 30 POTENTIAL.  
 FT CHAIN 31 254 KALLIKREIN 4.  
 FT ACT\_SITE 71 71 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 116 116 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 207 207 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 37 167 BY SIMILARITY.  
 FT DISULFID 56 72 BY SIMILARITY.  
 FT DISULFID 141 241 BY SIMILARITY.  
 FT DISULFID 148 213 BY SIMILARITY.  
 FT DISULFID 178 192 BY SIMILARITY.  
 FT DISULFID 203 228 BY SIMILARITY.  
 FT CARBOHYD 169 169 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 197 197 Q -> H (IN REF. 1 AND 4).  
 SQ SEQUENCE 254 AA; 27022 MW; 9C475E2B26EE0CB8 CR664;  
 Query Match 100.0%; Score 1369; DB 1; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-116;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATAGNFWGFLGLVILGVAGSLVSGSCQIINGEDCSHSPQWQALVWNEFLCSGVL 60  
 DB 1 MATAGNFWGFLGLVILGVAGSLVSGSCQIINGEDCSHSPQWQALVWNEFLCSGVL 60  
 QY 61 VHPQWVLSAHCQNSYITIGLHSLSEADQEPGQVMEASLSVRHPEYNRPFLANDLMLI 120  
 DB 61 VHPQWVLSAHCQNSYITIGLHSLSEADQEPGQVMEASLSVRHPEYNRPFLANDLMLI 120

QY 121 KLDESVSSEDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
 DB 121 KLDESVSSEDTIRISISIASQCPTAGNSCLVSGWGLLANGRMPTVLQCVNVSVSEVCSK 180  
 QY 181 LYDPLVHPNMFACGGQDQKDSGNSGGLICNGYLQGLVSGFKAPCGQGVGVYTNL 240  
 DB 181 LYDPLVHPNMFACGGQDQKDSGNSGGLICNGYLQGLVSGFKAPCGQGVGVYTNL 240  
 QY 241 CKFTEWIEKTVQAS 254  
 DB 241 CKFTEWIEKTVQAS 254  
 RESULT 2  
 ID KLK5 HUMAN STANDARD; PRT; 293 AA.  
 AC Q9Y377; Q9HBG8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Kallikrein 5 precursor (SC 3.4.21.-) (Stratum corneum tryptic enzyme)  
 DE (Kallikrein-like protein 2) (KLK-L2).  
 GN KLK5 OR SCTE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stratum corneum;  
 RX MEDLINE=99445563; PubMed=10514489;  
 RA Brattsand M., Egelrud T.;  
 RT "Purification, molecular cloning, and expression of a human stratum  
 RT corneum trypsin-like serine protease with possible function in  
 RT desquamation.";  
 RL J. Biol. Chem. 274:30033-30040(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20118156; PubMed=10652563;  
 RA Yousef G.M., Luo L.-Y., Diamandis E.P.;  
 RT "Identification of novel human kallikrein-like genes on chromosome  
 RT 19q13.3-q13.4.";  
 RL Anticancer Res. 19:2843-2852(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20510030; PubMed=11054574;  
 RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepel B., Wang K.;  
 RT "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RL Gene 257:119-130(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshivski S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumarane P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: May be involved in desquamation.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed in skin, breast, brain and testis.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF168768; AAF03101.1; -  
 DR EMBL; AF135028; AAD26429.1; -  
 DR EMBL; AF243527; AAG33358.1; -  
 DR EMBL; BC008036; AAH08036.1; -  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.017; -  
 DR Genew; HGNC:6366; KLK5.  
 DR MIN; 605643; -  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
 DR GO; GO:0008544; P:epidermal differentiation; TAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR HydroLase; Serine protease; Glycoprotein; Signal.  
 FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 293 KALLIKREIN 5.  
 FT ACT\_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 153 153 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 245 245 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 73 206 BY SIMILARITY.  
 FT DISULFID 93 109 BY SIMILARITY.  
 FT DISULFID 178 279 BY SIMILARITY.  
 FT DISULFID 185 251 BY SIMILARITY.  
 FT DISULFID 217 231 BY SIMILARITY.  
 FT DISULFID 241 266 BY SIMILARITY.  
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 25 56 MISSING (IN REF. 3).  
 SQ SEQUENCE 293 AA; 32020 MW; D92C92F5609E5946 CRC64;  
 Query Match 49.0%; Score 671; DB 1; Length 293;  
 Best Local Similarity 45.7%; Pred. No. 4.1e-53;  
 Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;  
 QY 1 MATAGNPNWFLGVI---LGV-----AGLSYG----- 26  
 DB 1 MATARPPMMWVLCALITALLGVTEHVLANDVSCDHPSTNVPFSGNQDLCAGAGEDARS 60  
 QY 27 --SCSQIINGDCSPHSQFQWQAALVME--NELFCGVLVHPQWVLSAAHCFQNSYITGLGL 83  
 DB 61 DSSSRILNGSDCDMHTQWQAALLRENLQYCGAVLVHPQWLLTAACRKKVFRVRLGH 120  
 QY 84 HSLADQFGQGVQWVLSVTHPEYNRPRLANDMLIKLDESVESTISIASOCP 143  
 DB 121 YSLSPVYESGQGMFGVKSIHPGYSHPGHNDMLIKLNRIRPTKDVRPINVSCHPS 180  
 QY 144 AGNSCLVSCWGLLANGRM--PTVLQCVNVSVVSEVSGSKLYDPLVHPMFCAAGGQDQKD 201

DB 181 AGTKCLVSGWGTTSQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFA-GDKAGRD 239  
 QY 202 SCNGSDGGPLICNGYLOGLVSGFGRAPCGQGVGVVYTNLCRFTWIEKTVQAS 254  
 DB 240 SCQGDGGPVVCGSLQGLVSGWGDYPCARPGRPVGVTNLCRFTWIOETIQAN 292  
 RESULT 3  
 NRPN RAT STANDARD; PRT; 260 AA.  
 AC 088780;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neuropein precursor (EC 3.4.21.-) (NP (Kallikrein 8) (Brain serine  
 DE protease 1).  
 GN KLK8 OR PRSS19 OR NRPN OR BSP1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=Fischer; TISSUE=Brain;  
 RX MEDLINE=98389725; PubMed=9722524;  
 RA Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathe R.;  
 RT "Serine proteases in rodent hippocampus."  
 RL J. Biol. Chem. 273:23004-23011(1998).  
 CC -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
 CC hippocampal plasticity. Has a strong proteolytic activity against  
 CC fibronectin (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.  
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
 CC -!- TISSUE SPECIFICITY: Restricted to hippocampus.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AJ005641; CAA06643.1; -  
 DR HSP; O61955; INPM.  
 DR MEROPS; S01.244; -  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW HydroLase; Serine protease; Glycoprotein; Zymogen; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT PROPEP 29 32 BY SIMILARITY.  
 FT CHAIN 33 260 NEUROPSIN.  
 FT ACT\_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 39 173 BY SIMILARITY.  
 FT DISULFID 58 74 BY SIMILARITY.  
 FT DISULFID 145 246 BY SIMILARITY.  
 FT DISULFID 152 218 BY SIMILARITY.  
 FT DISULFID 184 198 BY SIMILARITY.  
 FT DISULFID 208 233 BY SIMILARITY.  
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 260 AA; 28510 MW; 58DF4F0602A0B7F5 CRC64;  
 Query Match 41.6%; Score 569.5; DB 1; Length 260;



Db	131	VLPSRCEPPTCTVSGWGTTSPTDVTFFSPLMCDVVKLISPDQCTKVYKDLLENSMLC	190
Qy	193	AGGGQDKDCKNGSDGGLPLCNGLVQLVSGFKAPCGQGVGVYVNLCKFTWIKETVQ	252
Db	191	AGTPDSKKNACNGSDGGLPLCRGLTQGLVSGWTFPGQCPNDPGVYVQVCKFTKWIINDTKW	250
RESULT 5			
NRPN	MOUSE		
ID	NRPN	MOUSE	
AC	Q61955;	STANDARD;	PRT; 260 AA.
DT	15-JUL-1999	(Rel. 38, Created)	
DT	15-JUL-1999	(Rel. 38, Last sequence update)	
DT	15-MAR-2004	(Rel. 43, Last annotation update)	
DE	Neurosin precursor (EC 3.4.21.-) (NP) (Kallikrein 8).		
GN	KLK8 OR PRS19 OR NRPN		
OS	Mus musculus (Mouse)		
OC	Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/c; TISSUE=Hippocampus;		
RX	MEDLINE=95348817; PubMed=7623137;		
RA	Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,		
RA	Ito J., Nishino H., Amoto S., Kiyama H., Shiosaka S.;		
RT	"Expression and activity-dependent changes of a novel limbic-serine		
RT	protease gene in the hippocampus.";		
RL	J. Neurosci. 15:5088-5097(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Yoshida S., Hirata A., Inoue N., Shiosaka S.;		
RT	"Cloning and assignment of mouse neurosin gene, Prs19 to chromosome		
RT	7B4.";		
RL	Submitted (S2P-1999) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Czech II; TISSUE=Mammary gland;		
RX	MEDLINE=23388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,		
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RN	[4]		
RP	SEQUENCE OF N-TERMINUS, AND CHARACTERIZATION.		
RC	STRAIN=BALB/c; TISSUE=Brain;		
RX	MEDLINE=9825202; PubMed=9556608;		
RA	Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,		
RA	Shiosaka T., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;		
RT	"Characterization of recombinant and brain neurosin, a		
RT	plasticity-related serine protease.";		
RL	J. Biol. Chem. 273:11189-11196(1998).		
RN	[5]		
RP	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.		
RC	TISSUE=Hippocampus;		
RX	MEDLINE=99134351; PubMed=9933620;		
RA	Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,		

RA	Shiosaka S., Hakoshima T.;		
RT	"Crystal structure of neurosin, a hippocampal protease involved in		
RT	kindling epileptogenesis.";		
RL	J. Biol. Chem. 274:4220-4224(1999).		
CC	- FUNCTION: Suggested to be involved in kindling epileptogenesis and		
CC	hippocampal plasticity. Has a strong proteolytic activity against		
CC	fibronectin.		
CC	- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.		
CC	- ENZYME REGULATION: Strongly inhibited by diisopropyl		
CC	fluorophosphate, leupeptin and (4-aminophenyl)methanesulfonyl 1-		
CC	fluoride.		
CC	- SUBCELLULAR LOCATION: Secreted.		
CC	- TISSUE SPECIFICITY: Expressed specifically in the limbic system of		
CC	mouse brain and is localized at highest concentration in pyramidal		
CC	neurons of the hippocampal CA1-3 subfields.		
CC	- MASS SPECTROMETRY: MW=26613; METHOD=MALDI; RANGE=29-260.		
CC	- MASS SPECTROMETRY: MW=26229; METHOD=MALDI; RANGE=33-260.		
CC	- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; D30785; BAA06451.1; -		
DR	EMBL; AB032202; BAA92435.1; -		
DR	EMBL; BC055895; AAH55895.1; -		
DR	PIR; I56559; I56559.		
DR	PDB; INFM; 23-MAR-99.		
DR	MEROPS; S01.244; -		
DR	MGI; 892018; Klx8.		
DR	InterPro; IPR009003; Cys Ser trypsin.		
DR	InterPro; IPR01254; Peptidase S1.		
DR	InterPro; IPR001314; Peptidase_S1A.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	SMART; SM00020; Tryp_SPC; 1.		
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;		
KW	3D-structure. 1		
FT	SIGNAL	1	28
FT	PROPEP	29	32
FT	CHAIN	33	260
FT	ACT_SITE	73	73
FT	ACT_SITE	120	120
FT	ACT_SITE	212	212
FT	DISULFID	39	173
FT	DISULFID	58	74
FT	DISULFID	145	246
FT	DISULFID	152	218
FT	DISULFID	184	198
FT	DISULFID	208	233
FT	CARBOHYD	110	110
FT	STRAND	34	34
FT	STRAND	37	38
FT	TURN	41	42
FT	TURN	45	46
FT	TURN	47	52
FT	TURN	53	54
FT	TURN	55	64
FT	TURN	65	66
FT	TURN	72	70
FT	TURN	80	83
FT	TURN	87	87
FT	TURN	88	89
FT	TURN	96	98
FT	TURN	96	100
FT	STRAND	100	105
FT	STRAND	100	105

N-LINKED (GLCNAC. . .) (POTENTIAL).

FT	TURN	107	108
FT	TURN	114	115
FT	TURN	118	119
FT	STRAND	122	126
FT	STRAND	140	141
FT	TURN	148	149
FT	STRAND	151	156
FT	STRAND	170	170
FT	STRAND	172	178
FT	HELIX	181	187
FT	TURN	189	191
FT	TURN	194	195
FT	STRAND	196	200
FT	TURN	202	203
FT	STRAND	206	206
FT	TURN	209	210
FT	TURN	212	213
FT	STRAND	215	218
FT	TURN	219	220
FT	STRAND	221	228
FT	STRAND	235	235
FT	TURN	236	237
FT	STRAND	238	238
FT	STRAND	240	244
FT	HELIX	245	256
SQ	SEQUENCE	260 AA;	28523 MW; BE5F6F6BE37CD60E CRC64;

Query Match            41.2%; Score 563.5; DB 1; Length 260;  
 Best Local Similarity 41.3%; Pred.No.1.7e-43;  
 Matches 102; Conservative 47; Mismatches 91; Indels 7; Gaps 4;

QY	10	WFLGYLILGVAGSLVSGSCSIINGEDCSPHSQPWOAALWMENELFCSGVLVHPQWLVA	69
DB	12	WILLLLFWGAWGLTRAQGSKILEGRECIPIHSPQWOAALFQGERLICGGVLVGDRWLVLA	71
QY	70	AHCQNSVTIGLGHSLEADQEPGQSOMVEASLSVRHPEYNRL---	LANDLMLIKLDSEV 126
DB	72	AHCCKRKYSVELGDHSHSQSRDQP-EDEIQVAOSIQHPCTNNNPNPDHSDHIMLRQNQA	130
QY	127	SESOTISISIASOCPTAGNSCLIVSGMWGLLANGR--MPTVLOCVNVSVVSEVCSKLYDP	184
DB	131	NLGDYKEVQLANLCPKVGOKCIIISGWGITVTSQEINFNTLNCAEVKIYSQNKCEAYPG	190
QY	185	LYHPSMFAGGGQDKDSNGDSGGPLICNGYIQLGVISFGKAPCGQGVGPVGYINLKFT	244
DB	191	KITEGMVCA-GSSNGADTCQDSDGGPLVCMDMLQGITSWGSDPCGKPEKPGYVIKICRYT	249
QY	245	EWEIKTV 251	
DB	250	TWKXTM 256	

RESULT 6  
 KLBX HUMAN STANDARD; PRT; 250 AA.  
 ID KLBX HUMAN Q9UXK7; O75837; Q9NS65,  
 AC 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Kallikrein 11 precursor (EC 3.4.21.-) (Hippostasin) (Trypsin-like protease).  
 GN KLK11 OR PRSS20 OR TLSP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=98438738; PubMed=9765601;  
 RA Yoshida S., Taniguchi M., Suemoto T., Oka T., He X.P., Shiosaka S.; "cDNA cloning and expression of a novel serine protease, TLSP."; Biochim. Biophys. Acta 1399:225-228(1998).

[2] RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2).  
 RP TISSUE=Hippocampus, and Prostate;  
 RX MEDLINE=2032922; PubMed=10872828;  
 RA Mi-tsu S., Yamada T., Okui A., Komiani K., Uemura H., Yamaguchi N.;  
 RX "A novel isoform of a kallikrein-like protease, TLSP/hippocastin,  
 RT (PRSS20), is expressed in the human brain and prostate.";  
 RT Biochem. Biophys. Res. Commun. 272:205-211(2000).  
 [3] RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=20130117; PubMed=10652548;  
 RX Yousef G.M., Scorilas A., Diamandis E.P.;  
 RT "Genomic organization, mapping, tissue expression, and hormonal  
 RT regulation of trypsin-like serine protease (TLSP PRSS20), a new  
 RT member of the human kallikrein gene family.";  
 RT Genomics 63:88-96(2000).  
 [4] RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=20510030; PubMed=1054574;  
 RX Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
 RA Moss P., Paepel B., Wang K.;  
 RA "Sequencing and expression analysis of the serine protease gene  
 RT cluster located in chromosome 19q13 region.";  
 RT Gene 257:119-130(2000).  
 [5] RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
 RX Burkhardt-Schultz K., Gordon I., Dias J., Ramirez M., Scilwagen S.,  
 RA Phan H., Velasco N., Do L., Regala M., Terry A., Brower A., Garmes J.,  
 RA Danganan L., Exler A., Christensen M., Georgescu A., Avila J., Liu S.,  
 RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coeffield J.,  
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carraro A.V.;  
 RA "Sequence analysis of chromosome 19q13.4.";  
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 [6] RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Testis;  
 RX MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Tapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smillius D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RL -!- FUNCTION: Possible multifunctional protease. Efficiently cleaves  
 CC Bz-Phe-Arg-4-methylcoumaryl-7-amide, a kallikrein substrate, and  
 CC weakly cleaves other substrates for kallikrein and trypsin.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q9UBX7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q9UBX7-2; Sequence=VSP\_005402;  
 CC -!- TISSUE SPECIFICITY: Expressed in brain, skin and prostate. Isoform  
 CC 1 is expressed preferentially in brain; Isoform 2 in prostate.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB012917; BAA33404.1; ALT\_INIT.  
 EMBL; AB013730; BAA88713.1; -  
 EMBL; AB041036; BAA96797.1; -  
 EMBL; AF164623; AAD47815.1; -  
 EMBL; AF243527; AAG33364.1; -  
 EMBL; AC011473; AAG23257.1; -  
 EMBL; BC022668; AAX22068.1; -  
 HSSP; P00763; IDPO.  
 MEROPS; S01.257; -.  
 Genew; HGNC:6359; KLK11.  
 MW; 604434; -.  
 GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 InterPro; IPR009003; Cys Ser. trypsin.  
 InterPro; IPR001254; Peptidase\_S1.  
 InterPro; IPR001314; Peptidase\_S1A.  
 Pfam; PF00089; trypsin; 1.  
 PRINTS; PR00722; CHYMOTRYPSIN.  
 SMART; SMO020; Tryp\_Spc; 1.  
 PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 PROSITE; PS00135; TRYPSIN\_SER; 1.  
 Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen;  
 Alternative splicing.  
 SIGNAL 1 18  
 FT PROPEP 19 21 POTENTIAL.  
 FT CHAIN 22 250 ACTIVATION PEPTIDE (POTENTIAL).  
 FT ACT\_SITE 62 62 KALLIKREIN 11.  
 FT ACT\_SITE 110 110 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 28 163 BY SIMILARITY.  
 FT DISULFID 47 63 BY SIMILARITY.  
 FT DISULFID 135 237 BY SIMILARITY.  
 FT DISULFID 142 209 BY SIMILARITY.  
 FT DISULFID 174 188 BY SIMILARITY.  
 FT DISULFID 199 224 BY SIMILARITY.  
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 1 1 M -> MORLRGLDNKSSGRGLTAKEPGARSSPLQAM  
 (in isoform 2).  
 /FTID:VSP 005402.  
 SEQUENCE 250 AA; 27466 MW; 192D910BDCD7A56 CRC64;

Query Match 40.1%; Score 549; DB 1; Length 250;  
 Best Local Similarity 41.7%; Pred.No.3.3e-42;  
 Matches 103; Conservative 52; Mismatches 84; Indels 8; Gaps 4;

12 LGLYLTVAGSLVSGSCSQTINGEDCSPIHSQPQQAALVMENELFGSGVLHFQWVLSAAH 71  
 4 LQLILLALATGLVGGE-TRIIRKGFCKPHSQPQQAALFEKTRLLCGATLIAPRMLLTAAH 62  
 72 FQNSYTTGLGHLSEADQEPGQWVEASLVSRHPEYRNP-----LANDMLIKLSDSVS 127  
 63 CLKPRYIVHLQGNLQ-KEEGCEQTRTATESFPHPGFNNSLPNKDHRIIMLVKXASPVS 121  
 128 ESDTIRTSIASQCTAGNSCLVSGWGLLANG--RMPTVLQCVNVSVEEVCCKLYDPL 185  
 122 ITWAVEPTLSRVCVTAGTSLTISGWSGTSSPQLRPLPHTLRCANITIEHOKCENAYPGN 181  
 186 YHPSWFCAGGQDQKSDCNDSGGPLICNGYLQGLVSGFKAPCGQGVGVYTWLCKFTE 245  
 182 ITDTVMVCASVEGGKDSQCGSGPLVCNQLGGIIISWGODPCAIRTRKPGVTKVKRKVD 241

```
DR MEROPS; S01.029; -
DR Genew; HGNC:6362; KLK14.
DR MIM; 606135; -
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0005576; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Hydrolyase; Serine protease; Signal; Zymogen.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 24 ACTIVATION PEPTIDE (POTENTIAL).
FT ACT_SITE 25 251 KALLIKREIN 14.
FT ACT_SITE 67 67 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 111 111 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 204 204 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 31 164 BY SIMILARITY.
FT DISULFID 52 68 BY SIMILARITY.
FT DISULFID 143 210 BY SIMILARITY.
FT DISULFID 175 189 BY SIMILARITY.
FT DISULFID 200 225 BY SIMILARITY.
SQ SEQUENCE 251 AA; 27452 MW; 9087953BAFAVED25 CRC64;

Query Match 39.1%; Score 535.5; DB 1; Length 251;
Best Local Similarity 43.9%; Pred. No. 5.5e-41;
Matches 107; Conservative 34; Mismatches 92; Indels 11; Gaps 4;

Qy 16 ILGVAGSLVSGSCQINGEDCSFHSQWQAALVM--ENLFCGVLVHPQWLVAHCF 73
Db 10 VLAIAMTQSDENKIKGGTCTTSQFQWQAALLAGPRRFLCGGALLSGQWVITAAHCG 69

Qy 74 QNSVTIGLGLHSL--EADQPGQWYEASLSVRHPEYRPLLANDMLIKLDESYSSESD 130
Db 70 RPILQVALGXENLRWEATQ---QVLVRYQVTHPEYNSRTHDNDLMLQLQPPARIGR 125

Qy 131 TIRISIASOQPTAGNSCLVSGWGLLAN--GRMPTVLQCVNVSVSEVCSKLYDPLVHP 188
Db 126 AVRPIETVQACSPGTSRCVSGWGTTSPIARYPASLQCVNINISPDDEVQKAYPRITTP 185

Qy 189 SMFAGGQGDQKDCNGSGGGPLCNGVLQGLVSGFKAPQGVQVGVNLTCKFTSWIE 248
Db 186 GMVAGVPGGKDSQGGSGGFLVCRGQLGLVSGWGMERCALPGYGVYTNLCKYRSWIE 245

Qy 249 KTVQ 252
Db 246 ETMR 249

RESULT 8
KLKD HUMAN STANDARD; PRT; 277 AA.
ID KLKD HUMAN STANDARD; PRT; 277 AA.
AC Q9UKR3; Q9Y433;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)
DE (KLK-L4).
GN KLK13 OR KLK14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20229789; PubMed=10766816;
RA Yousef G.M., Chang A., Diamandis E.P.;
RT "Identification and characterization of KLK-L4, a new kallikrein-like
gene that appears to be down-regulated in breast cancer tissues.";
J. Biol. Chem. 275:11891-11898(2000).
[2]
SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
RA Burkhart-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Carnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Garrano A.V.;
RA "Sequence analysis of chromosome 19q13.4.";
RA Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 1-180 FROM N.A.
RC TISSUE=Uterus;
RA Ansoorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in prostate, breast, testis and
CC salivary gland.
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF135024; AAD26425.2; -
DR EMBL; AC011473; AAG23259.1; -
DR EMBL; AL050220; CAB43320.1; ALT_INIT.
DR HSP; P00763; IDPO.
DR MEROPS; S01.306; -
DR Genew; HGNC:6361; KLK13.
DR MIM; 605505; -
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0005508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Hydrolyase; Serine protease; Glycoprotein; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 277 KALLIKREIN 13.
FT ACT_SITE 76 76 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 124 124 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 218 218 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 178 BY SIMILARITY.
FT DISULFID 61 77 BY SIMILARITY.
FT DISULFID 157 224 BY SIMILARITY.
FT DISULFID 189 203 BY SIMILARITY.
FT DISULFID 214 239 BY SIMILARITY.
FT CARBOHYD 30 30 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 170 180 VNPFTLOCAN -> GMPHWPPEAP (IN REF. 3).
SQ SEQUENCE 277 AA; 30570 MW; BA8A98BDCFB5D542 CRC64;

Query Match 37.8%; Score 517.5; DB 1; Length 277;
Best Local Similarity 41.1%; Pred. No. 2.6e-39;
Matches 108; Conservative 46; Mismatches 90; Indels 19; Gaps 7;

Qy 7 PGWFLGVLILGVAGSLVSGSCQIN-----GDCSPHSQWQAALVMENELFC 56
```







```

FT DISULFID 184 198 BY SIMILARITY.
FT DISULFID 208 233 BY SIMILARITY.
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 23 A -> AAGSGDLTLKLYAENLPCVHLNPQWPSPQSHCPGRG
FT WSNPLPPAA (in isoform 2).
FT /FTId=VSP_005401.
SQ SEQUENCE 260 AA; 28048 MW; EF4395B8C3E660 CRC64;

Query Match 37.7%; Score 515.5; DB 1; Length 260;
Best Local Similarity 41.4%; Pred. No. 3.6e-39;
Matches 106; Conservative 39; Mismatches 96; Indels 15; Gaps 8;

QY 4 AGNPGWFLGVLILVAGSLVSGSCQ---IINGEDCSPHSQPQAALVWENELFCSGVL 60
DB 8 AAKTW-MFL--LLLG--GAWAGHRAQEDKVLGHGECQSPQRAALFQGLLGGVL 62
QY 61 VHPQWLGAACHFQNSYITIGLGLSLSEADQBPFGQMVSEASLVHPYNNRPLLA---NDL 117
DB 63 VGGNWLTAACHKPKYVRLGDHSLQKQGP-SQEIPVVGSIHPHCYNSSDVEDHNDL 121
QY 118 MLIKLDESVSDDTIRTSIASOCTAGNSCLVSGWGLLANGR--MPTVLQCNVSVSE 175
DB 122 MLLQLRQDASLGSKYKPSISLADHCTQPGQKCTVSGWGTVPSPRENFDTLNCARVKIPQ 181
QY 176 EVCSKLYDPLYPHSMFCAGGGQDQKSCNGSGGGLICNGYLQGLVFGKAPCQGVGPG 235
DB 182 KCEADAYEGQITDGMVCAAGSKG-ADTCQGSQGGPLVCDGALQGITSGSDPCGRSDKPG 240
QY 236 VYTNLCKFTWEIKTV 251
DB 241 VYTNICRYLDMIKKII 256

RESULT 10
ID KLKA RAT STANDARD; PRT; 244 AA.
AC P36375;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JAN-2004 (Rel. 43, Last annotation update)
DE Glandular kallikrein 10 precursor (EC 3.4.21.35) (Tissue kallikrein)
DE (T-kininogenase) (K10) (Proteinase B) (Endopeptidase K) (Fragment).
GN K10 OR KLK-10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE-Kidney, and Submandibular gland;
RX MEDLINE=93041794; PubMed=1420203;
RA Ma J.-X., Chao J., Chao L.;
RT "Molecular cloning and characterization of rKlk10, a cDNA encoding T-
RT kininogenase from rat submandibular gland and kidney.";
RL Biochemistry 31:10922-10928(1992).
RN [2]
SEQUENCE OF 10-32; 95-124 AND 179-232.
RP MEDLINE=91224135; PubMed=2026164;
RX Gutman N., Elmoujahed A., Brillard M., du Sorbier B., Gauthier F.;
RT "Microheterogeneity of rat submaxillary gland kallikrein k10, a
RT member of the kallikrein family.";
RL Eur. J. Biochem. 197:425-429(1991).
RN [3]
SEQUENCE OF 10-32 AND 97-133.
RP TISSUE=Submaxillary gland;
RX MEDLINE=90153911; PubMed=2303430;
RA Xiong W., Chen L.-M., Chao J.;
RT "Purification and characterization of a kallikrein-like
RT T-kininogenase.";
RL J. Biol. Chem. 265:2822-2827(1990).
RN [4]
SEQUENCE OF 10-32 AND 97-117.
RX MEDLINE=88198057; PubMed=3482210;

```

Query Match 37.6%; Score 515; DB 1; Length 244;  
 Best Local Similarity 42.4%; Pred. No. 3.7e-39;  
 Matches 101; Conservative 42; Mismatches 77; Indels 18; Gaps 5;

```

RA Kato H., Nakanishi E., Enryoji K., Hayaishi I., Oh-Ishi S., Iwanaga S.;
RT "Characterization of serine proteinases isolated from rat
RT submaxillary gland: with special reference to the degradation of rat
RT kininogens by these enzymes.";
RL J. Biochem. 102:1389-1404(1987).
CC -!- FUNCTION: Glandular kallikreins cleave Met-Lys and Arg-Ser bonds
CC in kininogen to release Lys-bradykinin. This protein may be
CC involved in the regulation of renal function.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage of Arg-|-Xaa bonds in
CC small molecule substrates. Highly selective action to release
CC kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of
CC Met-|-Xaa or Leu-|-Xaa.
CC -!- SUBUNIT: Heterodimer of a light chain and heavy chain linked
CC by a disulfide bond.
CC -!- TISSUE SPECIFICITY: Kidney and submandibular gland, where it is
CC found in the granular convoluted tubule and striated duct cells.
CC It is likely that the enzyme is mainly synthesized in the granular
CC convoluted tubules and then transferred to other tissues by
CC release into the vasculature or interstitial space.
CC -!- PTM: PROBABLY N- AND O-GLYCOSYLATED. IT HAS CARBOHYDRATE MOIETIES
CC OF ALPHA-METHYL-D-MANNOSIDE AND N-ACETYL-D-GLUCOSAMINE GROUPS.
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S48142; AAB24071.1; -
DR PIR; A44284; A44284.
DR PIR; B35545; B35545.
DR HSP; P00759; ITON.
DR MEROPS; S01.165;
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; TRYPTOPTRYSIN.
DR SMART; SM00020; TYR1_SPC; 1.
DR PROSITE; PS02240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HAS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR Hydrolase; Serine protease; Glycoprotein; Signal; Multigene family.
FT NON TER 1 1
FT SIGNAL <1 3 PROBABLE.
FT PROPEP 4 9 ACTIVATION PEPTIDE (PROBABLE).
FT CHAIN 10 244 GLANDULAR KALLIKREIN 10.
FT CHAIN 10 96 T-KININOGENASE LIGHT CHAIN.
FT CHAIN 97 244 T-KININOGENASE HEAVY CHAIN.
FT ACT_SITE 48 48 CHARGE RELAY SYSTEM.
FT ACT_SITE 103 103 CHARGE RELAY SYSTEM.
FT ACT_SITE 196 196 CHARGE RELAY SYSTEM.
FT DISULFID 15 156 BY SIMILARITY.
FT DISULFID 33 149 BY SIMILARITY.
FT DISULFID 135 202 BY SIMILARITY.
FT DISULFID 167 181 BY SIMILARITY.
FT DISULFID 192 217 BY SIMILARITY.
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 91 91 O-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 126 126 O-LINKED (POTENTIAL).
FT CARBOHYD 142 142 N -> IET (IN REF. 3).
FT CONFLICT 115 116 IT -> DS (IN REF. 3).
FT CONFLICT 128 128 E -> G (IN REF. 3).
FT CONFLICT 133 133 S -> G (IN REF. 3).
SQ SEQUENCE 244 AA; 27305 MW; BAB4D40547EB79C0 CRC64;

```

```
QY 29 SQIINGEDCSPHOPQAOALVMENELFCGVLVHPOMVLSAAHCFNSYTIIGLHSLA 88
DB 8 SRIVGKYCKEKNQPOVQWAI--NEVLCGGVLDPSPWITAAHCVSNYTHVLGRNLF 65
QY 89 DQFQSQMVASLSVRHPEYRPL-----ANDMLIKLDESVSDDTIRIS 136
DB 66 D-EPFAQYRFVNGSFPHDY-KPLFRNHTROGRDDYSNDMLLHLEPADITDGVKVID 123
QY 137 IASQCTAGNSCLVSGWLL--ANGEMPTVLOCVNVSVSEVCSKLYDPLVHPSMFCAG 194
DB 124 LPTPEPKVSTCLASGWSGSKPLNWLPELDLOQCNHLLSNKRCIAYEQKVTDLMLCAG 183
QY 195 GGDQDQSDNSGDSGGPLICNGYIQLGLVSFGKAPCGQGVGVGVVYTNLCCKTEWIEKTQV 252
DB 184 EMGRKDTCKGDSGGPLICDGVVQGITSGNVPFCAEPYNGVYTKLKFTSNKEVMK 241
```

```
RESULT 11
TRY3 CHICK STANDARD; PRT; 248 AA.
AC Q90629;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trypsin II-P29 precursor (EC 3.4.21.4).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=95251611; PubMed=7733885;
RA Wang K., Gan L., Lee L., Hood L.E.;
RT "Isolation and characterization of the chicken trypsinogen gene family.";
RL Biochem. J. 307:471-479(1995).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: High levels are seen in the pancreas while lower levels are found in the liver, spleen and thymus.
CC -1- SIMILARITY: Belongs to peptidase family S1.
```

```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; U15157; AAA79914.1; -
CC PIR; S55066; S55066.
CC HSP; P00763; 1DPO.
CC MEROPS; S01_151; -.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC K0 Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen;
CC Calcium-binding; Signal; Multigene family.
CC SIGNAL 1 16 BY SIMILARITY.
CC PROPEP 17 25 ACTIVATION PEPTIDE (BY SIMILARITY).
CC CHAIN 26 248 TRYPSIN II-P29.
CC ACT_SITE 65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT METAL 77 77 CALCIUM (BY SIMILARITY).
FT METAL 79 79 CALCIUM (VIA CARBONYL OXYGEN)
FT METAL 82 82 CALCIUM (BY SIMILARITY).
FT METAL 87 87 CALCIUM (VIA CARBONYL OXYGEN)
FT METAL 87 87 CALCIUM (BY SIMILARITY).
FT ACT SITE 109 109 CALCIUM (BY SIMILARITY).
FT CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 32 162 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 50 66 BY SIMILARITY.
FT DISULFID 134 235 BY SIMILARITY.
FT DISULFID 141 208 BY SIMILARITY.
FT DISULFID 173 187 BY SIMILARITY.
FT DISULFID 198 222 BY SIMILARITY.
FT SITE 196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 248 AA; 26622 MW; ESE16B07622B588E CRC64;
```

```
Query Match 37.3%; Score 510.5; DB 1; Length 248;
Best Local Similarity 41.5%; Pred. No. 9.6e-39;
Matches 102; Conservative 49; Mismatches 86; Indels 9; Gaps 6;
QY 14 YLILGVAGSLVS----GSCQIINGEDCSPHQPQAOALVMENELFCGVLVHPQWLSA 69
DB 5 FLILSCLGAAVAFPGGADDDKIVGGYTCPEHSPVQVSL-NSGYHFCGGSLSNQWLSA 63
QY 70 AHCFNSYTIIGLHSLAEOQPGSQMVASLSVRHPEYRPLANDMLIKLDESYSSES 129
DB 64 AHCKYSRIQVRLGEYNTDV-QEDSEVVRSSVIRHPKYSITLNNIMLIKASAVEYS 122
QY 130 DTIRISISTASOCTAGNSCLVSGWG-LLANG-RMPTVLQCVNVSVSEVCSKLYDPLH 187
DB 123 ADIOPALPSSCAKAGTECLISGNTLSNGYNYPELLQCLNAPILSDQCEAYPGDIT 182
QY 188 PSMFCAGGQDQKSCNGDSGGPLICNGYIQLGLVSFGKAPCGQGVGVGVVYTNLCCKTEW 247
DB 183 SNMICVGFLEGGKDCSQDGGPVVVCNGELQGIIVSWG-IGCALKGYPGVYTKVCNVYDVI 241
QY 248 EKTVOA 253
DB 242 QETIAA 247
```

```
RESULT 12
KLK3 MOUSE
ID KLK3 MOUSE STANDARD; PRT; 261 AA.
AC P00756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glandular kallikrein K3 precursor (EC 3.4.21.35) (Tissue kallikrein)
DE (mGK-3) (7S nerve growth factor gamma chain) (Gamma-NGF).
GN KLK3 OR KLK-3 OR NGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85076169; PubMed=6548955;
RX Ullrich A., Gray A., Wood W.I., Hayflick J., Seeburg P.H.;
RT "Isolation of a cDNA clone coding for the gamma-subunit of mouse nerve growth factor using a high-stringency selection procedure.";
RL DNA 3:387-392(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85257431; PubMed=3848399;
RX Evans B.A., Richards R.I.;
RT "Genes for the alpha and gamma subunits of mouse nerve growth factor are contiguous.";
RL EMBO J. 4:133-138(1985).
RN [3]
RP SEQUENCE OF 25-261.
RX MEDLINE=81264363; PubMed=7263706;
```

RA Thomas K.A., Baglan N.C., Bradshaw R.A.;  
RT "The amino acid sequence of the gamma-subunit of mouse submaxillary  
RT gland 7 S nerve growth factor.";  
RL J. Biol. Chem. 256:9156-9166(1981).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (3.15 ANGSTROMS) OF 7S COMPLEX.  
RC STRAIN-Swiss Webster; TISSUE-Submaxillary gland;  
RX MEDLINE=9803451; PubMed=9351801;  
RA Bax B., Blundell T.L., Murray-Rust J., McDonald N.Q.;  
RT "Structure of mouse 7S NGF: a complex of nerve growth factor with  
RT four binding proteins.";  
RL Structure 5:1275-1285(1997).  
CC -!- FUNCTION: 7S NGF alpha chain stabilizes the 7S complex. The beta  
CC dimer promotes neurite growth. The gamma chain is an arginine-  
CC specific protease; it may also have plasminogen activator  
CC activity as well as mitogenic activity for chick embryo  
CC fibroblasts.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage of Arg-|-Xaa bonds in  
CC small molecule substrates. Highly selective action to release  
CC kallidin (lysyl-bradykinin) from kininogen involves hydrolysis of  
CC Met-|-Xaa or Leu-|-Xaa.  
CC -!- COFACTOR: Binds 2 zinc ions per 7S complex. The zinc ions are  
CC bound at the alpha-gamma interfaces.  
CC -!- SUBUNIT: 7S nerve growth factor is composed of two alpha chains,  
CC a beta dimer composed of identical chains, and two gamma chains.  
CC -!- MISCELLANEOUS: This precursor is cleaved into segments to produce  
CC the active form of the gamma chain, which occurs naturally as  
CC combinations of either two or three segments held together by  
CC disulfide bonds: B1 and A, or B1, C and B2.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X01389; CAA25645.1; -  
CC EMBL; X01798; CAA25928.1; -  
CC EMBL; X01799; CAA25930.1; -  
CC PIR; A91005; NMGSG.  
CC PDB; 1SGF; 27-MAY-98.  
CC MEROPS; S01.170; -  
CC MGD; MGI:97322; Ngf.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SMC0020; TRYP\_SPC; 1.  
CC PROSITE; PS02440; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Hydrolase; Serine protease; Glycoprotein; Multigene family; Zymogen;  
KW Metal-binding; Zinc; Signal; Growth factor; 3D-structure.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 261  
FT CHAIN 25 107  
FT CHAIN 112 261  
FT ACT\_SITE 65 65  
FT ACT\_SITE 120 120  
FT ACT\_SITE 213 213  
FT METAL 231 231  
FT METAL 236 236  
FT DISULFID 31 173  
FT DISULFID 50 66  
FT DISULFID 152 219  
FT DISULFID 184 198  
FT DISULFID 209 234  
FT CARBOHYD 102 102  
FT N-LINKED (GLCNAC. . .).  
FT

FT DOMAIN 25 107  
FT DOMAIN 112 261  
FT DOMAIN 112 164  
FT DOMAIN 165 261  
FT CONFLICT 108 111  
FT STRAND 26 26  
FT STRAND 29 30  
FT STRAND 33 35  
FT TURN 37 38  
FT TURN 39 44  
FT TURN 45 46  
FT STRAND 47 56  
FT TURN 57 58  
FT STRAND 59 62  
FT HELIX 64 66  
FT STRAND 72 75  
FT STRAND 79 79  
FT TURN 80 81  
FT TURN 85 86  
FT TURN 88 97  
FT TURN 99 100  
FT HELIX 103 105  
FT TURN 118 119  
FT TURN 122 126  
FT TURN 148 149  
FT STRAND 151 156  
FT STRAND 159 159  
FT STRAND 167 167  
FT STRAND 170 170  
FT STRAND 172 179  
FT HELIX 181 187  
FT TURN 194 195  
FT STRAND 196 200  
FT STRAND 207 207  
FT TURN 210 211  
FT TURN 213 214  
FT STRAND 216 219  
FT TURN 220 221  
FT STRAND 222 229  
FT TURN 234 235  
FT TURN 237 238  
FT STRAND 241 245  
FT HELIX 246 249  
FT HELIX 250 258  
FT TURN 259 259  
SQ SEQUENCE 261 AA; 28998 MW; 4870748E174AF7C8 CRC64;  
Query Match 37.2%; Score 509; DB 1; Length 261;  
Best Local Similarity 39.5%; Pred.No. 1.4e-38;  
Matches 103; Conservative 48; Mismatches 86; Indels 24; Gaps 6;  
QY 10 WFL-GYLILGVAG-SLVSGSCSIINGEDCSPHSQPWQAALVMEELFCGVLVHPQWVL 67  
Db 2 WFLILFALLSLGGIDAAPPVQSGRIVGFKCKNSQPMHVAVRYTYQLCGVLLDENWVL 61  
QY 68 SAACFQNSYTTGLGLHSLEADQEPGSGQMVSEASLSVRHPEYRPL-AND 116  
Db 62 TAAHCYDDNYKWLGNLKNLFKD-EPSAQHRFVSKAIPHPGFNNSLMRKHIRELYDYSND 120  
QY 117 LMLIKLDESVSBSDTIRSIASQCPAGNSCLVSGWGLLANGRMPTVLQ-AND 170  
Db 121 LMLRLSKADITDVKPITLPTPEPKLSTGLASGWSIT---PTKQFTDDLVCVNL 176  
QY 171 SVVSEVCSKLYDPLVHPMFCAGGQDQKSCNGSDGGGLICNGVLOGVIFGKAPCGQ 230  
Db 177 KLLPNEDCAKAIKVTDMALCAGEMDGGKDTCKGDSGGGLICDGLVQIGTISWGTFCGE 236  
QY 231 VGVPGVYTNLCFTEWIEKTV 251  
Db 237 PDMFGVYTNLCNFTSWIKOTM 257  
RESULT 13

SEGMENT B1.  
SEGMENT A.  
SEGMENT C.  
SEGMENT B2.  
MISSING (IN REF. 2).





KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;  
KW Calcium-binding; Signal; Multigene family.  
FT SIGNAL 1 15  
FT PROPEP 16 24  
FT CHAIN 25 247  
FT ACT SITE 64 64  
FT METAL 76 76  
FT METAL 78 78  
FT METAL 81 81  
FT METAL 86 86  
FT METAL 108 108  
FT ACT SITE 201 201  
FT DISULFID 31 161  
FT DISULFID 49 85  
FT DISULFID 133 234  
FT DISULFID 140 207  
FT DISULFID 172 186  
FT DISULFID 197 221  
FT SITE 195 195  
SQ SEQUENCE 247 AA; 26269 MW; D74892BA584E48 CRC64;  
Query Match 37.0%; Score 506; DB 1; Length 247;  
Best Local Similarity 40.5%; Pred. No. 2.4e-38;  
Matches 100; Conservative 50; Mismatches 89; Indels 8; Gaps 6;  
QY 11 FLGYLILGVAGSL-VSGSCSQTIIINGDCSPHQPWQAALWMENELFCSGVLVHPQWLSA 69  
Db 6 FLAF--LGAVALPLDDDDKIVGGYTCQKNSLPYQVSL-NAGYHFCGSLINSQWVYSA 62  
QY 70 AHCFONSYYTIGLHSLADQEPGQMVASLSVRHPEYRRELLANDMLIKLDESVSSES 129  
Db 63 AHCKSRIOVRLGEHNIDV-VEGGEQFIDAAKTIHRPSYNANTFNDIMLIKINSPATLN 121  
QY 130 DTIRSIASQCTAGNSCLVSGWG--LLANGEMPTVLCQVNVSVSEVCSKLYDELYH 187  
Db 122 SRVSTVSLPRSCSGSGTKLVSGWNTLSGTYNPSLQCLDAPVLSDSCKSSYPGKIT 181  
QY 188 PSMFCAGGGQDQKSDSCNGDSGGPLICNGYLGVSFQKAPCGQGVGPVYTNLCXFTWI 247  
Db 182 SNMFCLGFLEGGKDSQGDGSGFPVNCVNGQLQGVVSWGYG-CAQKGPVYTKVCNVVNI 240  
QY 248 EKTQAS 254  
Db 241 QQTVAAN 247

Search completed: May 5, 2004, 16:08:25  
Job time : 18 secs